

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 11.5063 Seconds
(without alignments)
731.761 Million cell updates/sec

Title: US-10-083-336A-5

Perfect score: 1025

Sequence: 1 MIPPKYPIINFITAGATVQ.....ARQYIEGEMRIYNNRS 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/prodata/1/iaa/5A COMB.pcp.*

2: /cgn2_6/prodata/1/iaa/5B COMB.pcp.*

3: /cgn2_6/prodata/1/iaa/6A COMB.pcp.*

4: /cgn2_6/prodata/1/iaa/6B COMB.pcp.*

5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pcp.*

6: /cgn2_6/prodata/1/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1025	100.0	268	2	US-08-356-786-8
2	1025	100.0	534	2	US-08-356-786-10
3	1020	99.5	267	1	US-07-901-707-1
4	1020	99.5	267	1	US-07-988-430-1
5	1020	99.5	267	1	US-08-425-336-1
6	1020	99.5	267	1	US-08-488-113B-1
7	1020	99.5	267	1	US-08-477-484B-1
8	1020	99.5	267	2	US-08-646-360-1
9	1020	99.5	267	3	US-08-839-765-1
10	1020	99.5	267	3	US-09-136-389-1
11	1020	99.5	267	4	US-09-610-838-1
12	1020	99.5	267	5	PCT-US92-09487-1
13	1020	99.5	290	1	US-08-378-761A-27
14	1020	99.5	290	1	US-08-485-286-27
15	1020	99.5	290	6	5248606-4
16	1010	98.5	267	1	US-08-218-303-16
17	1010	98.5	267	2	US-08-338-793D-61
18	1010	98.5	267	4	US-09-538-873-1
19	930.5	90.8	540	1	US-08-378-761A-77
20	930.5	90.8	540	1	US-08-485-286-77
21	342	33.4	247	1	US-08-488-113B-6
22	342	33.4	247	1	US-08-477-484B-6
23	342	33.4	247	2	US-08-646-360-6
24	342	33.4	247	3	US-08-839-765-6
25	342	33.4	247	3	US-09-136-389-6
26	342	33.4	247	4	US-09-610-838-6
27	342	33.4	267	1	US-08-378-761A-74

28	342	33.4	267	1	US-08-485-286-74	Sequence 74, Appl
29	342	33.4	289	1	US-07-923-692C-4	Sequence 4, Appli
30	342	33.4	289	2	US-08-184-237-4	Sequence 4, Appli
31	342	33.4	289	2	US-08-482-920-4	Sequence 4, Appli
32	342	33.4	289	3	US-08-484-341-4	Sequence 4, Appli
33	342	33.4	289	3	US-08-483-502-4	Sequence 4, Appli
34	342	33.4	289	4	US-09-726-651A-4	Sequence 4, Appli
35	341.5	33.3	282	1	US-08-324-301-15	Sequence 15, Appl
36	329.5	32.1	250	1	US-08-378-761A-71	Sequence 71, Appl
37	329.5	32.1	250	1	US-08-485-286-71	Sequence 71, Appl
38	323.5	31.6	251	4	US-09-538-873-3	Sequence 3, Appli
39	312.5	30.5	255	1	US-07-901-707-6	Sequence 6, Appli
40	312.5	30.5	255	1	US-07-988-430-6	Sequence 6, Appli
41	312.5	30.5	255	1	US-08-425-336-6	Sequence 6, Appli
42	312.5	30.5	255	5	PCT-US92-09487-6	Sequence 6, Appli
43	312	30.4	248	3	US-08-902-486-7	Sequence 7, Appli
44	312	30.4	290	1	US-08-245-754A-2	Sequence 2, Appli
45	312	30.4	290	2	US-08-597-731-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-356-786-8
; Sequence 8, Application US/08356786
; Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS: 16
ADDRESSES: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-356-786-8

Query Match 100.0%; Score 1025; DB 2; Length 268;

Best Local Similarity 100.0%; Pred. No. 6e-112;

Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINQRFIL 60
DB 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINQRFIL 60
QY 61 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAP 120
DB 61 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAP 120
QY 121 GGNDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTARSFIICIMISEAA 180
DB 121 GGNDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTARSFIICIMISEAA 180
QY 181 RFOYIEGEMTRIRYNRRS 199
DB 181 RFOYIEGEMTRIRYNRRS 199

RESULT 2

US-08-356-786-10
; Sequence 10, Application US/08356786
; Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSER: Edmund R. Pitcher, Tests, Hurwitz, & Thibeault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,786
FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967
FILING DATE: 06-FEB-1992
ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-053
TELEPHONE: (617) 248-7000
TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 534 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-356-786-10

Query Match 100.0%; Score 1025; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 1.7e-111;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINQRFIL 60
DB 3 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINQRFIL 62
QY 61 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAP 120

DB 63 VELSNHAEISVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAP 122
QY 121 GGNDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTARSFIICIMISEAA 180
DB 123 GGNDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTARSFIICIMISEAA 182
QY 181 RFOYIEGEMTRIRYNRRS 199
DB 183 RFOYIEGEMTRIRYNRRS 201

RESULT 3

US-07-901-707-1
; Sequence 1, Application US/07901707
; Patent No. 5376546

GENERAL INFORMATION:

APPLICANT: Bernhardt, Susan L.
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Steve F.
APPLICANT: Lane, Julie A.
TITLE OF INVENTION: Materials Comprising and Methods of
TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
STREET: Two First National Plaza, 20 South Clark
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/901,707
FILING DATE: 19920619
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: No. 5376546and, Greta E.
REGISTRATION NUMBER: 35,302
REFERENCE/DOCKET NUMBER: 27129/30910

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 346-5750
TELEFAX: (312) 984-5750
TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 267 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein

US-07-901-707-1

Query Match 99.5%; Score 1020; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 2.3e-111;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IFFKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINQRFILV 61
DB 1 IFFKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNRVGLPINQRFILV 60

QY 62 ELSNHAELS VTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAP 121
DB 61 ELSNHAELS VTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAP 120
QY 122 GGNDRLEQLAGNLRNIELGNGLPLEEASALYYSTGGTQPLTARSFIICIMISEAA 181

Feb 15

Db 181 FOYIEGEMTRIRYNRS 198
 |||

RESULT 2

US-10-282-935-1
 ; Sequence 1, Application US/10282935
 ; Publication No. US20030143193A1

; GENERAL INFORMATION:

; APPLICANT: VITETTA, ELLEN S.

; APPLICANT: GHETIE, VICTOR F.

; APPLICANT: SMALLSHAW, JOAN

; APPLICANT: BALUNA, ROXANA G.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF

; TITLE OF INVENTION: PROTEINACEOUS COMPOUNDS

; FILE REFERENCE: UTSD:884US

; CURRENT APPLICATION NUMBER: US/10/282,935

; CURRENT FILING DATE: 2002-10-29

; PRIOR FILING DATE: 09/538,873

; PRIOR FILING DATE: 2000-03-30

; PRIOR APPLICATION NUMBER: 60/126,826

; PRIOR FILING DATE: 1999-03-30

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 267

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: Peptide

US-10-282-935-1

Query Match 100.0%; Score 1019; DB 12; Length 267;
 Best Local Similarity 100.0%; Pred. No. 2.2e-108;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILV 60

Db 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILV 60

Qy 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 120

Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 120

Qy 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFIIQIMISEAR 180

Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFIIQIMISEAR 180

Qy 181 FOYIEGEMTRIRYNRS 198

Db 181 FOYIEGEMTRIRYNRS 198

RESULT 3

US-10-440-796-1

; Sequence 1, Application US/10440796

; Publication No. US20040009148A1

; GENERAL INFORMATION:

; APPLICANT: VITETTA, ELLEN S.

; APPLICANT: GHETIE, VICTOR F.

; APPLICANT: SMALLSHAW, JOAN

; APPLICANT: BALUNA, ROXANA G.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK

; TITLE OF INVENTION: SYNDROME (VLS)

; FILE REFERENCE: UTSD:603

; CURRENT APPLICATION NUMBER: US/10/440,796

; CURRENT FILING DATE: 2003-05-19

; PRIOR FILING DATE: 09/538,873

; PRIOR FILING DATE: 2000-03-30

; PRIOR APPLICATION NUMBER: 60/126,826

; PRIOR FILING DATE: 1999-03-30

; NUMBER OF SEQ ID NOS: 19

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 267

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: Peptide

US-10-440-796-1

Query Match 100.0%; Score 1019; DB 12; Length 267;
 Best Local Similarity 100.0%; Pred. No. 2.2e-108;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILV 60

Db 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILV 60

Qy 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 120

Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 120

Qy 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFIIQIMISEAR 180

Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFIIQIMISEAR 180

Qy 181 FOYIEGEMTRIRYNRS 198

Db 181 FOYIEGEMTRIRYNRS 198

RESULT 4

US-10-083-336A-3

; Sequence 3, Application US/10083336A

; Publication No. US20030181665A1

; GENERAL INFORMATION:

; APPLICANT: Olson, Mark A

; APPLICANT: Millard, Charles B

; APPLICANT: Byrnie, Michael P

; APPLICANT: Wannenmacher, Robert W

; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof

; FILE REFERENCE: P67452US0 (RIID 01-58)

; CURRENT APPLICATION NUMBER: US/10/083,336A

; CURRENT FILING DATE: 2002-05-21

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 198

; TYPE: PRT

; ORGANISM: Ricinus communis

US-10-083-336A-3

Query Match 99.1%; Score 1010; DB 12; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.5e-107;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILV 62

Db 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILV 62

Qy 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 122

Db 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEAIHTLFTDVQNRVTFAG 122

Qy 123 YDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFIIQIMISEAR 182

Db 123 YDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQLTPLARSFIIQIMISEAR 182

Qy 183 YIEGEMTRIRYNRS 198

Db 183 YIEGEMTRIRYNRS 198

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 11.5641 Seconds
(without alignments)
731.761 Million cell updates/sec

Title: US-10-083-336A-10

Perfect score: 1029

Sequence: 1 MIFPKQYPIINFTAGATVQ.....RFQYIEGMRTRYNRSA 200

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1029	100.0	268	2	US-08-356-786-8
2	1029	100.0	534	2	US-08-356-786-10
3	1024	99.5	267	1	US-07-901-707-1
4	1024	99.5	267	1	US-07-988-430-1
5	1024	99.5	267	1	US-08-425-336-1
6	1024	99.5	267	1	US-08-488-113B-1
7	1024	99.5	267	1	US-08-477-484B-1
8	1024	99.5	267	2	US-08-546-360-1
9	1024	99.5	267	3	US-08-839-765-1
10	1024	99.5	267	3	US-09-136-389-1
11	1024	99.5	267	4	US-09-610-838-1
12	1024	99.5	267	5	PCT-US92-09487-1
13	1024	99.5	230	1	US-08-378-761A-27
14	1024	99.5	230	1	US-08-485-286-27
15	1024	99.5	230	6	5248606-4
16	1014	98.5	267	1	US-08-218-303-16
17	1014	98.5	267	2	US-08-338-793D-61
18	1014	98.5	267	4	US-09-538-873-1
19	934.5	90.8	540	1	US-08-378-761A-77
20	934.5	90.8	540	1	US-08-488-286-77
21	342	33.2	247	1	US-08-488-113B-6
22	342	33.2	247	1	US-08-477-484B-6
23	342	33.2	247	1	US-08-646-360-6
24	342	33.2	247	3	US-08-839-765-6
25	342	33.2	247	3	US-09-136-389-6
26	342	33.2	247	4	US-09-610-838-6
27	342	33.2	267	1	US-08-378-761A-74

28	342	33.2	267	1	US-08-485-286-74
29	342	33.2	289	1	US-07-923-692C-4
30	342	33.2	289	1	US-08-184-237-4
31	342	33.2	289	2	US-08-482-920-4
32	342	33.2	289	3	US-08-484-341-4
33	342	33.2	289	3	US-08-483-502-4
34	342	33.2	289	4	US-09-726-651A-4
35	341.5	33.2	282	1	US-08-324-301-15
36	329.5	32.0	250	1	US-08-378-761A-71
37	329.5	32.0	250	1	US-08-485-286-71
38	323.5	31.4	251	4	US-09-538-873-3
39	312.5	30.4	255	1	US-07-901-707-6
40	312.5	30.4	255	1	US-07-988-430-6
41	312.5	30.4	255	1	US-08-425-336-6
42	312.5	30.4	255	5	PCT-US92-09487-6
43	312	30.3	248	3	US-08-902-486-7
44	312	30.3	290	1	US-08-245-754A-2
45	312	30.3	290	2	US-08-597-731-2

ALIGNMENTS

RESULT 1
US-08-356-786-8
; Sequence B. Application US/08356786
; Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
TITLE OF INVENTION: Marker
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESS: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356.786
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/831,967

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Pitcher, Edmund R.

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: CRP-053

TELEPHONE:

TELEPHONE: (617) 248-7000

TELEFAX:

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO:

SEQ ID NO: 8

SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-356-786-8

Query Match 100.0%; Score 1029; DB 2; Length 268;
Best Local Similarity 100.0%; Pred. No. 2.2e-112;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 60
DB 1 MIFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 60
QY 61 VELSNAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAF 120
DB 61 VELSNAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAF 120
QY 121 GGYDRLEQLAGNLRNIELNGPLERAIISALYYSTGGTQLPTLARSFICIQMISEAA 180
DB 121 GGYDRLEQLAGNLRNIELNGPLERAIISALYYSTGGTQLPTLARSFICIQMISEAA 180
QY 181 RFOYIEGEMTRIRYNRRA 200
DB 181 RFOYIEGEMTRIRYNRRA 200

RESULT 2

US-08-356-786-10
; Sequence 10, Application US/08356786
; Patent No. 5677305

; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992

; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829

; REFERENCE/DOCKET NUMBER: CEP-053
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100

; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-08-356-786-10

Query Match 100.0%; Score 1029; DB 2; Length 534;
Best Local Similarity 100.0%; Pred. No. 6.1e-112;
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MIFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 60
DB 3 MIFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFIL 62
QY 61 VELSNAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAF 120

DB 63 VELSNAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAF 122
QY 121 GGYDRLEQLAGNLRNIELNGPLERAIISALYYSTGGTQLPTLARSFICIQMISEAA 180
DB 123 GGYDRLEQLAGNLRNIELNGPLERAIISALYYSTGGTQLPTLARSFICIQMISEAA 182
QY 181 RFOYIEGEMTRIRYNRRA 200
DB 183 RFOYIEGEMTRIRYNRRA 202

RESULT 3

US-07-901-707-1
; Sequence 1, Application US/07901707
; Patent No. 5376546
; GENERAL INFORMATION:

; APPLICANT: Bernhard, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Steve F.

; APPLICANT: Lane, Julie A.
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marchall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell,
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/901,707
; FILING DATE: 19920619
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991

; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5376546and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27129/30910

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-5750
; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear

; MOLECULE TYPE: protein
US-07-901-707-1

Query Match 99.5%; Score 1024; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 8.3e-112;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFILV 61
DB 1 IFPKQYPIINFTAGATVQSYTNFIRAVRGLTTGADVREHIEIPVLPNRVGLPINQRFILV 60
QY 62 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAF 121
DB 61 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDQVQRYTFAF 120
QY 122 GGYDRLEQLAGNLRNIELNGPLERAIISALYYSTGGTQLPTLARSFICIQMISEAA 181

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:35 ; Search time 10.7687 Seconds
(without alignments)
1777.145 Million cell updates/sec

Title: US-10-083-336A-5

Perfect score: 1025

Sequence: 1 MIFPKQPIINFITAGATVQ.....ARFQVIEGEMTRIRYNRRS 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020	99.5	576	1	RLCSD
2	930.5	90.8	564	1	RLCSAG
3	353.5	34.5	528	2	S32431
4	353.5	34.5	562	2	S16022
5	345	33.7	527	2	S32430
6	342	33.4	289	1	RLT2T
7	338	33.0	247	2	UT0393
8	338	33.0	247	2	UC5032
9	338	33.0	289	2	UC5606
10	336.5	32.8	251	2	C39761
11	329.5	32.1	528	1	TLZLSA
12	307.5	30.0	278	2	S23519
13	300.5	29.3	250	2	JN0108
14	291	28.4	570	2	S26227
15	286	27.9	254	2	P00108
16	284.5	27.8	277	2	S22494
17	279	27.2	286	2	S25560
18	278	27.1	245	1	RLP4840
19	277	27.0	286	1	RLPUGG
20	272	26.5	286	2	JC4235
21	242.5	23.7	316	2	JT0753
22	197.5	19.3	294	2	S28421
23	182	17.8	313	2	S17757
24	180.5	17.6	261	2	J50401
25	178	17.4	278	2	A39817
26	156	15.2	272	2	JC4811
27	149.5	14.6	289	2	T12573
28	140	13.7	280	1	RLBH
29	137	13.4	275	2	S33631

ALIGNMENTS

RESULT 1

RLCSD

ricin D precursor - castor bean

N:Contains: rRNA N-glycosidase (EC 3.2.2.22)

C:Species: Ricinus communis (castor bean)

C>Date: 31-Dec-1993 #sequence-revision 31-Dec-1993 #text change 16-Jul-1999

C:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903

R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.F.

Nucleic Acids Res 13, 8019-8033, 1985

A:Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.

A:Reference number: A24041; MUID:86067214; PMID:2999712

A:Accession: A24041

A:Molecule type: DNA

A:Residues: 1-576 <HAL>

A:Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083

R:Tregear, J.W.; Roberts, L.M.

Plant Mol. Biol. 18, 515-525, 1992

A:Title: The lectin gene family of Ricinus communis: cloning of a functional ricin gene

A:Reference number: S20513; MUID:92163016; PMID:1371405

A:Accession: S20513

A:Molecule type: DNA

A:Residues: 1-576 <TR>

A:Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085

R:Lamb, F.I.; Roberts, L.M.; Lord, J.M.

Eur. J. Biochem. 148, 265-270, 1985

A:Title: Nucleotide sequence of cloned cDNA coding for preprorizin.

A:Reference number: A24614; MUID:85179479; PMID:3838723

A:Accession: A24614

A:Molecule type: mRNA

A:Residues: 12-75, 'D', 77-550, 'R', 552-576 <LAM>

A:Cross-references: GB:X02388; NID:g21077; PIDN:CAA26230.1; PID:g21078

R:Yoshitake, S.; Funatsu, G.; Funatsu, M.

Agric. Biol. Chem. 42, 1267-1274, 1978

A:Title: Isolation and sequences of peptic peptides, and the complete sequence of Ile cha

A:Reference number: A03372

A:Accession: A03372

A:Molecule type: protein

A:Residues: 36-97, 'Q', 99-109, 'S', 111-269, 'D', 272-283, 'L', 285-288, 290-302 <YOS>

A:Note: this paper cites the others in the series providing experimental details for the

R:Araki, T.; Funatsu, G.

FEBS Lett. 191, 121-124, 1995

A:Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptophan

A:Reference number: A24010

A:Accession: A24010

A:Molecule type: protein

A:Residues: 315-383, 'PS', 386-576 <ARA>

R:Funatsu, G.; Kimura, M.; Funatsu, M.

Agric. Biol. Chem. 43, 2221-2224, 1979

A:Title: Primary structure of Ala chain of ricin D.

A:Reference number: A03374

A:Accession: A03374

A:Molecule type: protein

A:Residues: 315-335, 'N', 337-342, 'NH', 345-362, 364-383, 'PS', 386-399, 'T', 401, 'D', 403, 'E', 405, 'F', 407, 409-414, 'W', 566, 'H', 567-570, 'LI', 573-574, 'F' <FUN>
 A>Note: This paper, one of a series, summarizes the experimental details for the determination of the complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.
 R:Ready, M.P.; Kim, Y.; Robertus, J.D.
 Proteins 10, 270-278, 1991
 A:Title: Site-directed mutagenesis of ricin A-chain and implications for the mechanism of action of ricin.
 A:Reference number: A48237; PMID:91352006; PMID:1881883
 A:Contents: annotation; active site
 R:Rutenber, E.; Robertus, J.D.
 Proteins 10, 260-269, 1991
 A:Title: Structure of ricin B-chain at 2.5 angstrom resolution.
 A:Reference number: A48238; PMID:91352005; PMID:1881882
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms
 R:Katzin, B.J.; Collins, E.J.; Robertus, J.D.
 Proteins 10, 251-259, 1991
 A:Title: Structure of ricin A-chain at 2.5 angstroms.
 A:Reference number: A48239; PMID:91352004; PMID:1881881
 A:Contents: annotation; X-ray crystallography, 2.5 angstroms
 C:Comment: The functional molecule is a disulfide-linked dimer of A and B chains, which is inactivated by protein synthesis; it inactivates the 60S ribosomal subunit into the cell of the A chain; B chains are also responsible for cell agglutination (lectin).
 C:Comment: This protein is cytotoxic and very poisonous to animals.
 C:Superfamily: ricin; rRNA N-glycosidase homology
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed
 F:1-35/Domain: signal sequence #status predicted <SIG>
 F:36-302/Product: ricin D chain A #status experimental <ACH>
 F:46-293/Domain: rRNA N-glycosidase homology <RNG>
 F:315-376/Product: ricin D chain B #status experimental <BCH>
 F:331-373, 374-414, 417-455, 462-497, 501-540, 543-576/Region: 40-residue repeats
 F:45, 409, 449/Binding site: carboxylate (Asn) (covalent) #status experimental
 F:115, 158, 243, 244/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F:212/Active site: Glu #status experimental
 F:215/Active site: Arg #status predicted
 F:294-318, 334-353, 377-394, 465-478, 504-521/Disulfide bonds: #status experimental
 F:336, 349, 360/Binding site: N-acetylglactosamine (Asp, Glu, Asn) #status experimental
 F:548, 569/Binding site: N-acetylglactosamine (Asp, Asn) #status experimental

Query Match 99.5%; Score 1020; DB 1; Length 576;
 Best Local Similarity 100.0%; Pred. No. 58-84; Indels 0; Gaps 0;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IPFKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNVRVGLPINQRFILV 61
 |||||
 Db 36 IPFKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNVRVGLPINQRFILV 95

QY 62 ELSNHAELSVTLLDVTNAYVVGGRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVYTFAG 121
 |||||
 Db 96 ELSNHAELSVTLLDVTNAYVVGGRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVYTFAG 155

QY 122 GNYDRLEQLAGNRLNIELGNGPLEAISALYYSTGGTQPLTARSFFIICQMISEAAR 181
 |||||
 Db 156 GNYDRLEQLAGNRLNIELGNGPLEAISALYYSTGGTQPLTARSFFIICQMISEAAR 215

QY 182 FOYIEGEMTRIRYNRRS 199
 |||||
 Db 216 FOYIEGEMTRIRYNRRS 233

RESULT 2
 RLCSAG
 agglutinin precursor - castor bean
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
 C:Species: Ricinus communis (castor bean)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C:Accession: A24261; A24210
 R:Robertus, L.M.; Lamb, I.F.; Pappin, D.J.C.; Lord, J.M.
 J. Biol. Chem. 260, 15682-15686, 1985
 A:Title: The primary sequence of Ricinus communis agglutinin. Comparison with ricin.
 A:Reference number: A24261; PMID:86059449; PMID:2999130
 A:Accession: A24261
 A:Molecule type: mRNA
 A:Residues: 1-564 <ROB>
 A:Cross-references: GB:M12089; NID:g169700; PIDN:AAA33869.1; PID:g169701

R:Araki, T.; Yoshioka, Y.; Funatsu, G.
 Biochim. Biophys. Acta 872, 277-285, 1986
 A:Title: The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin.
 A:Reference number: A24210
 A:Accession: A24210
 A:Molecule type: Protein
 A:Residues: 303-325, 'F', 327-330, 'T', 332-361, 'D', 363-373, 'G', 375-403, 'T', 405-551, 'V', 553-554, 'E', 556-566, 'H', 567-570, 'LI', 573-574, 'F' <FUN>
 A:Comment: This protein has strong agglutinating activity and weak cytotoxicity compared with ricin.
 C:Superfamily: ricin; rRNA N-glycosidase homology
 C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; RNA binding; seed
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-290/Product: agglutinin chain A #status predicted <ACH>
 F:303-564/Product: agglutinin chain B #status experimental <BCH>
 F:319-361, 362-402, 405-443, 450-485, 489-528, 531-564/Region: 40-residue repeats
 F:34, 259/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:104, 147, 231, 232/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
 F:200, 203/Active site: Glu, Arg #status predicted
 F:282-306, 322-341, 365-382, 453-466, 492-509/Disulfide bonds: #status predicted
 F:324, 337, 348/Binding site: N-acetylglactosamine (Asp, Glu, Asn) #status predicted
 F:397, 437/Binding site: carboxylate (Asn) (covalent) #status experimental
 F:536, 557/Binding site: N-acetylglactosamine (Asp, Asn) #status predicted

Query Match 90.8%; Score 930.5; DB 1; Length 564;
 Best Local Similarity 91.9%; Pred. No. 5.6e-76; Indels 1; Gaps 1;
 Matches 182; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 2 IPFKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNVRVGLPINQRFILV 61
 |||||
 Db 25 IPFKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNVRVGLPINQRFILV 84

QY 62 ELSNHAELSVTLLDVTNAYVVGGRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVYTFAG 121
 |||||
 Db 85 ELSNHAELSVTLLDVTNAYVVGGRAGNSAYFFHPDQEDAEAIHTLFTDVQNRVYTFAG 144

QY 122 GNYDRLEQLAGNRLNIELGNGPLEAISALYYSTGGTQPLTARSFFIICQMISEAAR 181
 |||||
 Db 145 GNYDRLEQLAGNRLNIELGNGPLEAISALYYSTGGTQPLTARSFFIICQMISEAAR 203

QY 182 FOYIEGEMTRIRYNRRS 199
 |||||
 Db 204 FOYIEGEMTRIRYNRRS 221

RESULT 3
 S32431
 abrin-d precursor - Indian licorice (fragment)
 N:Contains: rRNA N-glycosidase (EC 3.2.2.22)
 C:Species: Abrus precatorius (Indian licorice)
 C:Date: 30-Sep-1993 #sequence_revision 01-Aug-1997 #text_change 01-Aug-1997
 C:Accession: S32431; S34408
 R:Hung, C.H.; Lee, M.C.; Lee, T.C.; Lin, J.Y.
 J. Mol. Biol. 229, 263-267, 1993
 A:Title: Primary structure of three distinct isoabrin determinants determined by cDNA sequencing.
 A:Reference number: S32429; PMID:93132798; PMID:8421313
 A:Accession: S32431
 A:Molecule type: mRNA
 A:Residues: 1-528 <HUN>
 A:Cross-references: GB:M98346
 R:Hung, C.; Lee, M.; Lee, T.; Lin, J.
 submitted to the EMBL Data Library, March 1993
 A:Reference number: S34408
 A:Accession: S34408
 A:Molecule type: mRNA
 A:Residues: 1-169, 'C', 171-320, 'L', 322-528 <HU2>
 A:Cross-references: GB:M98346
 C:Comment: Abrin consists of an A chain, which inhibits protein synthesis by inactivating the A and B chains are linked by a single disulfide bond, which is essential for toxicity.
 C:Superfamily: ricin; rRNA N-glycosidase homology
 C:Keywords: disulfide bond; duplication; glycoprotein; glycosidase; hydrolase; lectin; p
 F:1-251/Product: abrin-d chain A #status predicted <ACH>
 F:7-246/Domain: rRNA N-glycosidase homology <RNG>
 F:261-528/Product: abrin-d chain B #status predicted <BCH>


```

Best Local Similarity   39.5%;   Pred.No. 2.6e-23;
Matches    73; Conservative   46; Mismatches    54; Indels     12; Gaps      5;

QY      10 INFTTAGATVQSYTFNIRAVGRLLTGADVRIEIPVLPNRVGLPINORFTILVELSNHAEL 69
       ::|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB      25 VSFRLSGATSSYGVFISNLKALPNERKL-YDIPILL--RSLPLGSQRVALIHLTNVADE 81

QY      70 SVTLALDVNTNAVVGVRAGNSAYFFHPDNQEDA-EAITHLFTDVQNRYTFAFGNYDRLE 128
       ::|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB      82 TTSVAIDVTNVIMGYRADGTGYFF---NEASATEAAKVFDKAERKVTLPYSGNVERLQ 138

QY      129 QLAGNLRNIELNGNPLEPAISALYYYSYGGTQLPTLARSFIIICIQWISEAARFOYIEGE 188
       ::|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB      139 TAAGKIRENIPLGLPALDSAITTLFYNNAN----SAASALMWLIQTSEAAKYKTEIQ 193

QY      189 METRI 193
       :|:
DB      194 IGRKV 198

RESULT 7
JU0393
karasurin - Mongolian snake-gourd
C:Species: Trichosanthes kirilowii (Mongolian snake-gourd)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 12-Apr-1995
C:Accession: JU0393; PS0163
R:Toyokawa, S.; Takeda, T.; Kato, Y.; Wakabayashi, K.; Ogiwara, Y.
Chem. Pharm. Bull. 39, 1244-1249, 1991
A:Title: The complete amino acid sequence of an abortifacient protein, karasurin
A:Reference number: JU0393; PMID:92005921; PMID:1914000
A:Accession: JU0393
A:Molecule type: protein
A:Residues: 1-247 <TOY>
A>Note: a sequence which lacks Ala-247 is also shown in this publication
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
C:KeywordDB: abortifacient
F:4-243/Domain: rRNA N-glycosidase homology <RNG>

Query Match          33.0%; Score 338; DB 2; Length 247;
Best Local Similarity 40.0%; Pred. No. 4.9e-23;
Matches    74; Conservative   46; Mismatches    53; Indels     12; Gaps      5;

QY      10 INFTTAGATVQSYTFNIRAVGRLLTGADVRIEIPVLPNRVGLPINORFTILVELSNHAEL 69
       ::|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB      2  VSFRLSGATSSYGVFISNLKALPYERKL-YDIPILL--RSTPLPGSQRYVALIHLTNVADE 58

QY      70 SVTLALDVNTNAVVGVRAGNSAYFFHPDNQEDA-EAITHLFTDVQNRYTFAFGNYDRLE 128
       ::|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB      59 TTSVAIDVTNVIMGYRADGTGYFF---NEASATEAAKVFDKAERKVTLPYSGNVERLQ 115

QY      129 QLAGNLRNIELNGNPLEPAISALYYYSYGGTQLPTLARSFIIICIQWISEAARFOYIEGE 188
       ::|::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB      116 IAAGKIRENIPLGLPALDSAITTLFYNNAN----SAASALMWLIQTSEAAKYKTEIQ 170

QY      189 METRI 193
       :|:
DB      171 IGRKV 175

RESULT 8
JC5032
karasurin-B - Trichosanthes kirilowii var. japonica
C:Species: Trichosanthes kirilowii var. japonica
C>Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 23-May-1997
C:Accession: JC5032
R:Kondo, T.; Mizukami, H.; Takeda, T.; Ogiwara, Y.
Biол. Pharm. Bull. 19, 1485-1489, 1996
A:Title: Amino acid sequences and ribosome-inactivating activities of karasurin
A:Reference number: JC5032; PMID:97108848; PMID:89511169
A:Accession: JC5032
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-247 <KON>
```


A;Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein from
 A;Reference number: JTO202
 A;Accession: JTO202
 A;Molecule type: protein
 A;Residues: 1-201, 203-251 <FUN>
 A;Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
 R;Evensen, G.; Mathiesen, A.; Sundan, A.
 J. Biol. Chem. 266, 6848-6852, 1991
 A;Title: Direct molecular cloning and expression of two distinct abrin A-chains.
 A;Reference number: A39761; MUID:91201329; PMID:2016300
 A;Accession: A39761
 A;Status: nucleic acid sequence not shown
 A;Molecule type: DNA
 A;Residues: 'E', 2-251 <EVE>
 A;Cross-references: GB:X54872
 A;Note: residues 1-8 were derived from the synthesized primer
 R;Kimura, M.; Sumizawa, T.; Funatsu, G.
 Biosci. Biotechnol. Biochem. 57, 166-169, 1993
 A;Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic
 A;Reference number: JCI398; MUID:93169023; PMID:7763422
 A;Contents: seeds
 A;Accession: JCI398
 A;Molecule type: protein
 A;Residues: 261-347, 'T', 349-351, 'A', 353-357, 'L', 359-528 <KIM>
 A;Experimental source: seed
 R;Evensen, G.; Mathiesen, A.; Sundan, A.
 submitted to the EMBL Data Library, October 1990
 A;Description: Direct molecular cloning of two distinct abrin A-chains.
 A;Reference number: S14471
 A;Accession: S14472
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 'ME', 2-251 <EV2>
 A;Cross-references: EMBL:X54873; NID:g16090; PIDN:CAA38655.1; PID:g16091
 R;Chen, Y.H.; Chow, L.P.; Tsugita, A.; Lin, J.Y.
 FEBS Lett. 309, 115-118, 1992
 A;Title: The complete primary structure of abrin-a B chain.
 A;Reference number: S24133; MUID:92371656; PMID:1505674
 A;Accession: S24133
 A;Molecule type: protein
 A;Residues: 262-297, 'Y', 299-426, 'L', 428-466, 'P', 468-482, 'L', 484-528 <CHE>
 R;Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.
 Eur. J. Biochem. 240, 564-569, 1996
 A;Title: Probing the domain structure of abrin-a by tryptic digestion.
 A;Reference number: S74110; MUID:97008945; PMID:8856055
 A;Accession: S74110
 A;Molecule type: protein
 A;Residues: 89-108, 154-172 <LIN>
 A;Experimental source: seed
 A;Accession: S74111
 A;Molecule type: protein
 A;Residues: 262-276, 'X', 278-280, 329-348, 369-388, 399-418 <LIW>
 A;Experimental source: seed
 C;Comment: Abrin-a is more toxic than ricin. The toxin consists of an A chain, which inh
 taining receptors on the cell surface. The A and B chains are linked by a single disulf
 C;Superfamily: ricin; rRNA N-glycosidase homology
 C;Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglutamic acid
 F;1-251/Product: abrin-a chain A #status experimental <ACH>
 F;7-246/Domain: rRNA N-glycosidase homology <RNG>
 F;261-528/Product: abrin-a chain B #status experimental <BCH>
 F;283-325, 326-366, 369-407, 414-449, 453-492, 495-528/Region: 40-residue repeats
 F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
 F;74, 113, 195, 196/Binding site: substrate (Tyr, Glu, Asn) #status predicted
 F;164, 167/Active site: Glu, Arg #status predicted
 F;247-269, 286-305, 329-346, 417-430, 456-473/Disulfide bonds: #status predicted
 F;288, 312/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted
 F;361, 401/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F;500, 521/Binding site: N-acetylgalactosamine (Asp, Asn) #status predicted
 Query Match 32.1%; Score 329.5; DB 1; Length 528;
 Best Local Similarity 42.8%; Pred. NO. 7.5e-22;
 Matches 80; Conservative 28; Mismatches 66; Indels 13; Gaps 5;

QY 10 INFTTAGATVQSYTNFIRAVRGLTITGADVRHEIPVLPNVRGLPINQRFILVELSNHAEL 69
 DB 5 IKFSTEGATSYKQFIEALRRLGG--LIHDIPLVLPDPTTLQERNRYITVELSNSOTE 62
 QY 70 SVTLALDVNTAVVGYRAGNSAYFFH--PDNQDAEALTHLFTDVQNRYYTFAFGNYDLR 127
 DB 63 SIEVGIDVTNAVYVAYRAGTSYFLRDAPSSASD-----YLFTGT-DQHSLPFYGYGDL 116
 QY 128 EQLAGNLENIELNGPLLEAISAIIYYSTGTQTPTLARSPFIICIMISEARFOYIEG 187
 DB 117 ERWAHQSROQIPLGHALTHGIS--FFRSGGNDNEEKARTLIVIQWAAEAREFYISN 173
 QY 188 EMRTIR 194
 DB 174 RVRVSIQ 180
 RESULT 12
 S23519
 beta-luffin - smooth loofah
 C;Species: Luffa cylindrica (smooth loofah)
 C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Aug-1999
 C;Accession: S23519; S23113
 R;Kataoka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Koizumi, A.
 Plant Mol. Biol. 19, 887-889, 1992
 A;Title: Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-inactivating
 A;Reference number: S23519; MUID:92353400; PMID:1643290
 A;Accession: S23519
 A;Molecule type: mRNA
 A;Residues: 1-278 <KAT>
 A;Cross-references: EMBL:X62372; NID:g19149; PIDN:CAA44230.1; PID:g19150
 C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
 F;26-264/Domain: rRNA N-glycosidase homology <RNG>
 Query Match 30.0%; Score 307.5; DB 2; Length 278;
 Best Local Similarity 35.8%; Pred. NO. 3.2e-20;
 Matches 67; Conservative 45; Mismatches 64; Indels 11; Gaps 3;
 QY 10 INFTTAGATVQSYTNFIRAVRGLTITGADVRHEIPVLPNVRGLPINQRFILVELSNHAEL 69
 DB 24 VPSLUGADSKSYKFFITALKALPSKEKVNIPLLPSASGA---SRVILMQLSNYDAK 80
 QY 70 SVTLALDVNTAVVGYRAGNSAYFFH--PDNQDAEALTHLFTDVQNRYYTFAFGNYDLR 129
 DB 81 AITMAIDVTNVYINGLVNSTSYFF--NESDAKLASQYVFKGSTIVTLPYSGNYERLQN 137
 QY 130 LAGNLENIELNGPLLEAISAIIYYSTGTQTPTLARSPFIICIMISEARFOYIEGEM 189
 DB 138 AAGKREKIPLGPRAFDSAITSLFHYDS-----TAAAGAFVLIIQTAAASREKYEQI 192
 QY 190 RTRIRYN 196
 DB 193 IERIPKN 199
 RESULT 13
 JN0108
 luffin-b - smooth loofah
 C;Species: Luffa cylindrica (smooth loofah)
 C;Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 07-May-1999
 C;Accession: JN0108
 R;Islam, M.R.; Hirayama, H.; Funatsu, G.
 Agric. Biol. Chem. 55, 229-238, 1991
 A;Title: Complete amino acid sequence of luffin-b, a ribosome-inactivating protein from
 A;Reference number: JN0108; MUID:91248488; PMID:1368666
 A;Accession: JN0108
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 1-250 <ISL>
 C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
 F;5-246/Domain: rRNA N-glycosidase homology <RNG>
 Query Match 29.3%; Score 300.5; DB 2; Length 250;

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004. CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:55 ; Search time 6.63825 Seconds
(without alignments)
1409.756 Million cell updates/sec

Title: US-10-083-336A-5

Perfect score: 1025

Sequence: 1 MIFPKQYPIINFETAGATVQ.....ARFQYIEGEMTRIRYNRRS 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1020	99.5	576	1 RICI_RICCO	P02879 ricinus com
2	930.5	90.8	564	1 AGEL_RICCO	P06750 ricinus com
3	353.5	34.5	562	1 ABRP_ABRPR	P28590 abrus preca
4	345	33.7	527	1 ABRB_ABRPR	Q06077 abrus preca
5	342	33.4	289	1 RIPT_TRIKI	P09989 trichosanthe
6	341.5	33.3	282	1 RIPT_BRYDI	P98184 bryonia dio
7	338	33.0	289	1 RIPS_TRIKI	P24478 trichosanthe
8	330.5	32.2	563	1 NIGB_SAMNI	P33183 sambucus ni
9	329.5	32.1	528	1 ABRP_ABRPR	P11140 abrus preca
10	312	30.4	290	1 RIPI_BRYDI	P33185 bryonia dio
11	300.5	29.3	250	1 RIPI_LUFY	P22851 luffa cylin
12	296.5	28.9	286	1 RIPI_CUCFI	Q9194 cucumis fig
13	286	27.9	254	1 MLA_VISAL	P81446 viscum albu
14	284.5	27.8	277	1 RIPI_LUFY	Q00465 luffa cylin
15	279	27.2	286	1 RIPI_MOMCH	P29339 momordica b
16	277	27.0	286	1 RIPI_MOMCH	P16094 momordica c
17	260	25.4	294	1 RIPI_TRIAN	P56626 trichosanthe
18	242.5	23.7	316	1 RIPI_GELMU	P33186 gelonium mu
19	197.5	19.3	294	1 RIPI_PHYAM	Q03464 phytoacca
20	182	17.8	313	1 RIPI_PHYAM	P10297 phytoacca
21	180.5	17.6	261	1 RIPS_PHYAM	P23339 phytoacca
22	175	17.4	278	1 RIPI_MIRJA	P21326 mirabilis j
23	140	13.7	280	1 RIPI_HORVU	P04399 hordeum vul
24	133	13.0	280	1 RIPI_HORVU	P22444 hordeum vul
25	131	12.8	253	1 RIPI_SAPOF	Q41391 saponaria o
26	127	12.4	253	1 RIPI_SAPOF	Q41389 saponaria o
27	126	12.3	310	1 RIPI_PHYAM	Q40772 phytoacca
28	124	12.1	299	1 RIPI_SAPOF	P20656 saponaria o
29	123.5	12.0	293	1 RIPI_DIACA	P24476 dianthus ca
30	121	11.8	292	1 RIPI_SAPOF	P27559 saponaria o
31	111	10.8	319	1 SLTA_BP933	P09385 bacterioph
32	110	10.7	236	1 RIPI_SAPOF	P27560 saponaria o
33	109.5	10.7	300	1 RIPI_MAIZE	P25891 zea mays (m

34	107.5	10.5	301	1 RIPX_MAIZE	P28522 zea mays (m
35	103	10.0	304	1 RIP9_MAIZE	P25892 zea mays (m
36	102	10.0	560	1 JI60_HORVU	Q00531 hordeum vul
37	91	8.9	315	1 SLTA_BPH19	P08026 bacterioph
38	91	8.9	315	1 SLTA_BPH30	P10149 bacterioph
39	85.5	8.3	1165	1 SVV_AQUAE	O67411 aquifex aeo
40	85	8.3	609	1 FTSH_TREPA	O83746 treponema p
41	82.5	8.0	492	1 AERA_AERTR	P09166 aeromonas t
42	81	7.9	305	1 Y348_MYCPN	P75205 mycoplasma
43	78.5	7.7	485	1 AERS_AERHY	Q06306 aeromonas h
44	78.5	7.7	493	1 AERA_AERHY	P09167 aeromonas h
45	76	7.4	157	1 RIP4_SAPOF	P27561 saponaria o

ALIGNMENTS

RESULT 1
ID RICI_RICCO STANDARD; PRT; 576 AA.
AC P02879; P02880;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ricin precursor [Contains: Ricin A chain (rRNA N-glycosidase)
DE (EC 3.2.2.22); Ricin B chain].
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067214; PubMed=2999712;
RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,
RA Weaver R.F.;
RT "Genomic cloning and characterization of a ricin gene from Ricinus
communis";
RL Nucleic Acids Res. 13:8019-8033(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92163016; PubMed=1371405;
RA Tregear J.W., Roberts L.M.;
RT "The lectin gene family of Ricinus communis: cloning of a functional
ricin gene and three lectin pseudogenes.";
RL Plant Mol. Biol. 18:515-525(1992).
RN [3]
RP SEQUENCE OF 12-576 FROM N.A.
RX MEDLINE=85179479; PubMed=3838723;
RA Lamb A., Roberts L.M., Lord J.M.;
RT "Nucleotide sequence of cloned cDNA coding for preproricin.";
RL Eur. J. Biochem. 148:265-270(1985).
RN [4]
RP SEQUENCE OF 36-302.
RX Yoshitake S., Funatsu G., Funatsu M.;
RT "Isolation and sequences of peptic peptides, and the complete
sequence of the chain of ricin-D.";
RL Agric. Biol. Chem. 42:1267-1274(1978).
RN [5]
RP SEQUENCE OF 315-576.
RX Funatsu G., Kimura M., Funatsu M.;
RT "Primary structure of Ala chain of ricin D.";
RL Agric. Biol. Chem. 43:2221-2224(1979).
RN [6]
RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.
RX MEDLINE=90344223; PubMed=1368517;
RA Kimura Y., Kusuoku H., Tada M., Takagi S., Funatsu G.;
RT "Structural analyses of sugar chains from ricin A-chain variant.";
RL Agric. Biol. Chem. 54:157-162(1990).
RN [7]
RP REVIEW.
RX MEDLINE=21480122; PubMed=11595634;
RA Olmes S., Kozlov J.V.;

RT "Ricin."; CC
 RL Toxicol 39:1723-1728 (2001). CC
 RN [8] CC
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS). CC
 RX MEDLINE=87165983; PubMed=3558397; CC
 RA Montfort W., Villafraña J.E., Monzingo A.F., Ernst S.R., Katzin B., CC
 RA Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.; CC
 RT "The three-dimensional structure of ricin at 2.8 Å."; CC
 RL J. Biol. Chem. 262:5398-5403 (1987). CC
 RN [9] CC
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN. CC
 RX MEDLINE=91352004; PubMed=1881881; CC
 RA Katzin B.J., Collins E.J., Robertus J.D.; CC
 RT "Structure of ricin A-chain at 2.5 Å."; CC
 RL Proteins 10:251-259 (1991). CC
 RN [10] CC
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN. CC
 RX MEDLINE=91352005; PubMed=1881882; CC
 RA Rutenber E., Robertus J.D.; CC
 RT "Structure of ricin B-chain at 2.5-Å resolution."; CC
 RL Proteins 10:260-269 (1991). CC
 RN [11] CC
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN. CC
 RX MEDLINE=95082010; PubMed=7990130; CC
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J., CC
 RA Faupcit R.A.; CC
 RT "X-ray structure of recombinant ricin A-chain at 1.8-Å resolution."; CC
 RL J. Mol. Biol. 244:410-422 (1994). CC
 RN [12] CC
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215. CC
 RX MEDLINE=96374222; PubMed=8780513; CC
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M., CC
 RA Molina-Svinth M.C., Robertus J.D.; CC
 RT "Structure and activity of an active site substitution of ricin A CC
 RT chain."; CC
 RL Biochemistry 35:11098-11103 (1996). CC
 RN [13] CC
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN. CC
 RX MEDLINE=97240820; PubMed=9086280; CC
 RA Yan X., Hollis T., Svinth M., Day P., Monzingo A.F., Milne G.W., CC
 RA Robertus J.D.; CC
 RT "Structure-based identification of a ricin inhibitor."; CC
 RL J. Mol. Biol. 266:1043-1049 (1997). CC
 RN [14] CC
 RP MUTAGENESIS. CC
 RX MEDLINE=93165632; PubMed=1287657; CC
 RA Kin Y., Robertus J.D.; CC
 RT "Analysis of several key active site residues of ricin A chain by CC
 RT mutagenesis and X-ray crystallography."; CC
 RL Protein Eng. 5:775-779 (1992). CC
 CC -!- FUNCTION: Ricin is highly toxic to animal cells and to a less CC
 CC extent to plant cells. The A chain is responsible for inhibiting CC
 CC protein synthesis through the catalytic inactivation of 60S CC
 CC ribosomal subunits. It acts as a glycosidase that removes a CC
 CC specific adenine residue from an exposed loop of 28S ribosomal CC
 CC RNA. As this loop is involved in the binding of elongation CC
 CC factors, the modified ribosomes are unable to support protein CC
 CC synthesis. The A chain can inactivate a few thousand ribosomes CC
 CC per minute, thus inactivating them faster than the cell can make CC
 CC new ones. A single A-chain molecule can therefore kill an animal CC
 CC cell. The B chain binds to cell receptors and facilitates the CC
 CC entry into the cell of the A chain; B chains are also responsible CC
 CC for cell agglutination (lectin activity). It binds to beta-D- CC
 CC galactopyranoside moieties. CC
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one CC
 CC specific adenosine on the 28S rRNA. CC
 CC -!- SUBUNIT: Disulfide-linked dimer of A and B chains. CC
 CC -!- DOMAIN: The B chain is composed of two domains, each domain CC
 CC consists of 3 homologous subdomains (alpha, beta, gamma). CC
 CC -!- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE CC
 CC MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271. CC
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME- CC
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY. CC

CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains. CC
 CC -!- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS CC
 CC WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3). CC
 CC -!- DATABASE: NAME=Protein Spotlight; CC
 CC NOTE=Issue 31 of February 2003; CC
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt031.html". CC
 CC ----- CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation CC
 CC the European Bioinformatics Institute. There are no restrictions on its CC
 CC use by non-profit institutions as long as its content is in no way CC
 CC modified and this statement is not removed. Usage by and for commercial CC
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ CC
 CC or send an email to license@isb-sib.ch). CC
 CC ----- CC
 CC EMBL; X03179; CAA26939.1; -; CC
 CC EMBL; X52908; CAA37095.1; -; CC
 CC EMBL; X02388; CAA26230.1; -; CC
 CC EMBL; A12892; CAA01058.1; -; CC
 CC FIR; A24041; RLCSD. CC
 CC PDB; 2AAI; 31-JAN-94. CC
 CC PDB; 1AFG; 31-JAN-94. CC
 CC PDB; 1FMP; 31-OCT-93. CC
 CC PDB; 1IFS; 14-JAN-98. CC
 CC PDB; 1IFT; 14-JAN-98. CC
 CC PDB; 1IFU; 14-JAN-98. CC
 CC PDB; 1FTC; 31-OCT-93. CC
 CC PDB; 1OBS; 16-JUN-97. CC
 CC PDB; 1OBT; 16-JUN-97. CC
 CC PDB; 1BR5; 02-SEP-98. CC
 CC PDB; 1BR6; 02-SEP-98. CC
 CC PDB; 1IL3; 16-JAN-02. CC
 CC PDB; 1IL4; 16-JAN-02. CC
 CC PDB; 1IL9; 16-JAN-02. CC
 CC GlycoSuiteDB; P02879; -; CC
 CC InterPro; IPR000772; Ricin_B_lectin. CC
 CC InterPro; IPR001574; RIP. CC
 CC Pfam; PF00652; Ricin_B_lectin; 6. CC
 CC Pfam; PF00161; RIP; 1. CC
 CC PRINTS; PR00396; SHIGARICIN. CC
 CC SMART; SM00458; RICIN; 2. CC
 CC PROSITE; PS02331; RICIN_B_LECTIN; 2. CC
 CC PROSITE; PS00275; SHIGA_RICIN; 1. CC
 CC KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; CC
 CC Glycoprotein; Lectin; Signal; 3D-structure. CC
 CC SIGNAL 1 35 CC
 CC FT CHAIN 36 302 CC
 CC FT PEPTIDE 303 314 CC
 CC FT CHAIN 315 576 CC
 CC FT DOMAIN 321 448 CC
 CC FT DOMAIN 451 575 CC
 CC FT REPEAT 331 373 CC
 CC FT REPEAT 374 414 CC
 CC FT REPEAT 417 449 CC
 CC FT REPEAT 462 497 CC
 CC FT REPEAT 501 540 CC
 CC FT REPEAT 543 570 CC
 CC FT ACT_SITE 212 212 CC
 CC FT DISULFID 294 318 CC
 CC FT DISULFID 334 353 CC
 CC FT DISULFID 377 394 CC
 CC FT DISULFID 465 478 CC
 CC FT DISULFID 504 521 CC
 CC FT CARBOHYD 45 45 CC
 CC FT CARBOHYD 271 271 CC
 CC FT CARBOHYD 409 409 CC
 CC FT CARBOHYD 449 449 CC
 CC FT CARBOHYD 76 76 CC
 CC FT CONFLICT 551 551 CC
 CC FT STRAND 43 47 CC
 CC FT TURN 49 50 CC

N-LINKED (GLCNAC. . .)
 /FTIG-CAR 000080.
 N-LINKED (GLCNAC. . .) (IN MINOR FORM).
 /FTIG-CAR 000081.
 N-LINKED (GLCNAC. . .).
 N-LINKED (GLCNAC. . .).
 E -> D (IN REF. 3).
 A -> R (IN REF. 3).

Query Match 99.5%; Score 1020; DB 1; Length 576;
 Best Local Similarity 100.0%; Pred. No. 2.7e-87;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREIPLVPLNRVGLPINQRFILV 61
 DB 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREIPLVPLNRVGLPINQRFILV 95

QY 62 ELSNHAELSVTLALDVTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRITFAFG 121
 DB 96 ELSNHAELSVTLALDVTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRITFAFG 155

QY 122 GNYDRLEQLAGNLRNLELNGPLNEEASALYYSTGCTQPLTARSFICQMISEAAR 181
 DB 156 GNYDRLEQLAGNLRNLELNGPLNEEASALYYSTGCTQPLTARSFICQMISEAAR 215

QY 182 FOYIEGEMRTIRYNRRS 199
 DB 216 FOYIEGEMRTIRYNRRS 233

RESULT 2

AGGL RICCO STANDARD; PRT; 564 AA.

AC P06750;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-glycosidase) (EC 3.2.2.22); Agglutinin B chain].
 DE Glycosidase (EC 3.2.2.22); Agglutinin B chain].
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
 OX NCBI_TaxID=3988;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86059449; PubMed=2999130;
 RA Roberts L.M., Lamb F.I., Pappin D.J.C., Lord J.M.;
 RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin.";
 RL J. Biol. Chem. 260:15682-15686(1985).
 RN [2]
 RP SEQUENCE OF 303-564.

RC TISSUE=Seed;
 RA Araki T., Yoshioka Y., Funatsu G.;
 RT "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";
 RL Biochim. Biophys. Acta 872:277-285(1986).
 RN [3]
 RP SEQUENCE OF 303-337.
 RX MEDLINE=80178723; PubMed=6768555;
 RA Lin T.-S., Li S.-L.;
 RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis.";
 RL Eur. J. Biochem. 105:453-459(1980).
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DB EMBL; M12089; AAA33869.1; -
 DB EMBL; S40368; AAB22584.1; -

DR PIR; A24261; RLCSAG.
 DR HSSP; P02879; IIR6.
 DR GlycosylatedB; P06750; -
 DR InterPro; IPR000772; Ricin_B_lectin.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00652; Ricin_B_lectin; 6.
 DR Pfam; PF00161; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS00231; RICIN B LECTIN; 2.
 DR PROSITE; PS00275; SHIGA RICIN; 1.
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 KW Glycoprotein; Lectin; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 290 AGGLUTININ A CHAIN.
 FT PROPEP 291 302 LINKER PEPTIDE.
 FT CHAIN 303 564 AGGLUTININ B CHAIN.
 FT DOMAIN 309 436 RICIN B-TYPE LECTIN 1.
 FT DOMAIN 439 563 RICIN B-TYPE LECTIN 2.
 FT REPEAT 319 361 1-ALPHA.
 FT REPEAT 362 402 1-BETA.
 FT REPEAT 405 437 1-GAMMA.
 FT REPEAT 450 485 2-ALPHA.
 FT REPEAT 489 528 2-BETA.
 FT REPEAT 531 558 2-GAMMA.
 FT ACT SITE 200 200 BY SIMILARITY.
 FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 322 341 BY SIMILARITY.
 FT DISULFID 365 382 BY SIMILARITY.
 FT DISULFID 453 466 BY SIMILARITY.
 FT DISULFID 492 509 BY SIMILARITY.
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 331 331 F -> T (IN REF. 2).
 FT CONFLICT 362 362 N -> D (IN REF. 2).
 FT CONFLICT 374 374 R -> G (IN REF. 2).
 FT CONFLICT 404 404 R -> T (IN REF. 2).
 FT CONFLICT 552 552 F -> V (IN REF. 2).
 SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CRC64;

Query Match 90.8%; Score 930.5; DB 1; Length 564;
 Best Local Similarity 91.9%; Pred. No. 5.6e-79;
 Matches 182; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREIPLVPLNRVGLPINQRFILV 61
 DB 25 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVREIPLVPLNRVGLPINQRFILV 84

QY 62 ELSNHAELSVTLALDVTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRITFAFG 121
 DB 85 ELSNHAELSVTLALDVTNAVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRITFAFG 144

QY 122 GNYDRLEQLAGNLRNLELNGPLNEEASALYYSTGCTQPLTARSFICQMISEAAR 181
 DB 145 GNYDRLEQLAGNLRNLELNGPLNEEASALYYSTGCTQPLTARSFICQMISEAAR 203

QY 182 FOYIEGEMRTIRYNRRS 199
 DB 204 FOYIEGEMRTIRYNRRS 221

RESULT 3

ID ABRCA ABRPR STANDARD; PRT; 562 AA.
 AC P28590;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Abrin-c precursor [Contains: Abrin-c A chain (rRNA N-glycosidase) (EC 3.2.2.22); Abrin-c B chain].
 OS Abrus precatorius (Indian licorice) (Crab's eye).

Query Match 34.5%; Score 353.5; DB 1; Length 562;

```

DR EMBL; M98345; AAA32625.1; -.
DR PIR; S32430; S32430.
DR HSSP; P11140; LABR.
DR InterPro; IPR000772; Ricin_B_lectin.
DR DR InterPro; IPR001574; RIP.
DR DR Pfam; PF00652; Ricin_B_lectin; 6.
DR DR Pfam; PF00161; RIP; 1.
DR DR PRINTS; PR00396; SHIGARICIN.
DR DR SMART; SM00458; RICIN; 2.
DR DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
DR DR PROSITE; PS0275; SHIGA_RICIN; 1.
KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
KW Glycoprotein; Lectin; Pyrrolidone carboxylic acid.
FT CHAIN 1 250 ABRIN-B A CHAIN.
FT PEPTIDE 251 260 LINKER PEPTIDE.
FT CHAIN 261 527 ABRIN-B B CHAIN.
FT DOMAIN 261 527 ABRIN-B B CHAIN.
FT DOMAIN 272 399 RICIN B-TYPE LECTIN 1.
FT DOMAIN 402 526 RICIN B-TYPE LECTIN 2.
FT REPEAT 282 324 1-ALPHA.
FT REPEAT 325 365 1-BETA.
FT REPEAT 368 400 1-GAMMA.
FT REPEAT 413 448 2-ALPHA.
FT REPEAT 452 491 2-BETA.
FT REPEAT 494 527 2-GAMMA.
FT ACT_SITE 163 163 BY SIMILARITY.
FT DISULFID 246 268 INTERCHAIN (BY SIMILARITY).
FT DISULFID 285 304 BY SIMILARITY.
FT DISULFID 328 345 BY SIMILARITY.
FT DISULFID 416 429 BY SIMILARITY.
FT DISULFID 455 472 BY SIMILARITY.
FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY).
FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 282 282 N -> D (IN REF. 2).
FT CONFLICT 291 291 D -> N (IN REF. 2).
FT CONFLICT 350 351 AE -> PO (IN REF. 2).
FT CONFLICT 378 378 S -> N (IN REF. 2).
FT CONFLICT 426 426 L -> M (IN REF. 2).
FT CONFLICT 428 428 Y -> D (IN REF. 2).
FT CONFLICT 431 431 N -> S (IN REF. 2).
FT CONFLICT 484 484 R -> K (IN REF. 2).
FT CONFLICT 491 491 N -> S (IN REF. 2).
FT CONFLICT 493 493 H -> Y (IN REF. 2).
FT CONFLICT 502 502 R -> G (IN REF. 2).
FT CONFLICT 509 509 E -> Q (IN REF. 2).
FT CONFLICT 513 513 H -> W (IN REF. 2).
FT CONFLICT 516 516 H -> T (IN REF. 2).
SQ SEQUENCE 527 AA; 59114 MW; 3253AE490CE9494A CRC64;
Query Match 33.7%; Score 345; DB 1; Length 527;
Best Local Similarity 45.5%; Pred. No. 1.5e-24;
Matches 87; Conservative 22; Mismatches 72; Indels 10; Gaps 4;
QY 6 OYPIINFTAGATVOSYTNFIRAVRGRLTGADVRHEIPVLPNRCVLPINQIFVLVLSN 65
Db 1 QDQIKFTTEGATSQSYKQFIEALQRUTGG--LIHGIPVLPDPTTLQERNYIIVLSN 58
QY 66 HAELSVTLALDVNTAYVGVYRAGNSAYFFPHDPNQEDAEATHLFTDVQNRVYTFAPFGNYD 125
Db 59 SDTESIEAGIDVSNAYVAVYRAGNSYFL--RDAPTSASRYLFTCTQ-QYSLRFGSYI 114
QY 126 RLEQLAGNIRENIELNGNPLEIAISALYYVYSGTGQTLPLARSFICIMISEARFQVI 185
Db 115 DLERLARQTRQIPLGLQALRHAISFL----QSGTDDQEIARTLIVIIQWASEAARYREI 170
QY 186 EGEEMTRIRYN 196
Db 171 SYRVGVGSIRTN 181

```

RESULT 5

```

RIPT_TRIKI
ID RIPT_TRIKI STANDARD; PRT; 289 AA.
AC P09989;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ribosome-inactivating protein alpha-trichosanthin precursor
DE (RNA N-glycosidase) (EC 3.2.2.22) (Alpha-TCS).
OS Trichosanthes kirilowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxID=3677;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz;
RA MEDLINE=91153657; PubMed=199291;
RX Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;
RT "Cloning of trichosanthin cDNA and its expression in Escherichia coli.";
RL Gene 97:267-272 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Maximowicz; TISSUE=Leaf;
RA MEDLINE=90256790; PubMed=2341400;
RX Chow T., Feldman R.A., Lovett M., Piatak M.;
RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a type I ribosome-inactivating protein.";
RL J. Biol. Chem. 265:8670-8674 (1990).
RN [3]
RP SEQUENCE OF 24-270.
RC STRAIN=Maximowicz; TISSUE=Tuberous root;
RA MEDLINE=90256789; PubMed=2341399;
RX Collins E.J., Robertus J.D., Iopresti M., Stone K.L., Williams K.R.,
RT Wu P., Hwang K., Piatak M.;
RT "Primary amino acid sequence of alpha-trichosanthin and molecular models for abrin A-chain and alpha-trichosanthin.";
RL J. Biol. Chem. 265:8665-8669 (1990).
RN [4]
RP SEQUENCE OF 24-270.
RC TISSUE=Tuberous root;
RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,
RT Tian G.Y., Ni C.Z.;
RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and application.";
RL Pure Appl. Chem. 58:789-798 (1986).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).
RA MEDLINE=94344957; PubMed=8066085;
RX Zhou F., Fu Z., Chen M., Lin Y., Pan K.;
RT "Structure of trichosanthin at 1.88-A resolution.";
RL Proteins 19:4-13 (1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RA MEDLINE=95344383; PubMed=7619070;
RX Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
RT depurinating mechanism of two ribosome-inactivating proteins.";
RL Biochem. J. 309:285-298 (1995).
CC -!- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS
CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT
CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to license@isb-sib.ch.

```

CC CC EMBL; M34858; AAA34207.1; -.
CC DR EMBL; J05434; AAA34206.1; -.
CC DR PIR; J0566; K1Z1.
CC DR PDB; 1MRJ; 07-FEB-95.
CC DR PDB; 1MRK; 07-FEB-95.
CC DR PDB; 1TCS; 10-JUL-95.
CC DR PDB; 1J4G; 28-JAN-03.
CC DR PDB; 1NLI; 21-JAN-03.
CC DR PDB; 1QD2; 24-APR-00.
CC DR InterPro; IPR001574; RIP.
CC DR Pfam; PF00161; RIP; 1.
CC DR PRINTS; PR00396; SHIGARICIN.
CC DR PROSITE; PS00275; SHIGA RICIN; 1.
CC DR Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
CC KW Toxin; Signal; 3D-structure.
CC FT SIGNAL 1 23
CC FT CHAIN 24 270
CC FT
CC FT PROPEP 271 289
CC FT ACT SITE 183 183
CC FT CONFLICT 57 60
CC FT CONFLICT 82 84
CC FT CONFLICT 87 87
CC FT CONFLICT 92 92
CC FT CONFLICT 92 92
CC FT CONFLICT 143 144
CC FT CONFLICT 196 196
CC FT CONFLICT 215 216
CC FT CONFLICT 231 231
CC FT CONFLICT 234 234
CC FT CONFLICT 246 266
CC FT CONFLICT 247 247
CC FT STRAND 25 28
CC FT TURN 30 31
CC FT HELIX 34 46
CC FT TURN 47 47
CC FT STRAND 50 54
CC FT TURN 55 56
CC FT STRAND 57 60
CC FT HELIX 66 69
CC FT STRAND 70 76
CC FT TURN 78 79
CC FT STRAND 82 88
CC FT TURN 89 92
CC FT STRAND 93 99
CC FT TURN 100 101
CC FT STRAND 102 105
CC FT HELIX 109 114
CC FT TURN 115 117
CC FT TURN 120 121
CC FT STRAND 124 127
CC FT HELIX 134 141
CC FT TURN 142 142
CC FT HELIX 145 147
CC FT STRAND 150 150
CC FT HELIX 152 163
CC FT TURN 164 165
CC FT HELIX 167 180
CC FT TURN 181 181
CC FT HELIX 182 186
CC FT STRAND 187 187
CC FT HELIX 188 195
CC FT TURN 196 196
CC FT STRAND 202 202
CC FT HELIX 206 226
CC FT TURN 227 230
CC FT STRAND 231 239
CC FT TURN 241 242
CC FT STRAND 245 250
CC FT TURN 251 252
CC FT HELIX 254 258
CC FT TURN 259 259
CC FT
CC FT MISSING IN MATURE PROTEIN.
CC FT BY SIMILARITY.
CC FT IPLL -> LPLI (IN REF. 4).
CC FT MISSING (IN REF. 4).
CC FT I -> L (IN REF. 4).
CC FT V -> VDAGLPNAV (IN REF. 4).
CC FT KI -> GL (IN REF. 4).
CC FT K -> S (IN REF. 4).
CC FT WS -> LWL (IN REF. 4).
CC FT Q -> T (IN REF. 4).
CC FT S -> T (IN REF. 2).
CC FT MISSING (IN REF. 4).
CC FT T -> M (IN REF. 2).
CC FT
CC FT RIBOSOME-INACTIVATING PROTEIN ALPHA-
CC FT TRICHOSANTHIN.
CC FT
CC FT MISSING IN MATURE PROTEIN.
CC FT BY SIMILARITY.
CC FT IPLL -> LPLI (IN REF. 4).
CC FT MISSING (IN REF. 4).
CC FT I -> L (IN REF. 4).
CC FT V -> VDAGLPNAV (IN REF. 4).
CC FT KI -> GL (IN REF. 4).
CC FT K -> S (IN REF. 4).
CC FT WS -> LWL (IN REF. 4).
CC FT Q -> T (IN REF. 4).
CC FT S -> T (IN REF. 2).
CC FT MISSING (IN REF. 4).
CC FT T -> M (IN REF. 2).
CC FT
CC FT RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
CC FT PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).
CC FT CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC FT specific adenosine on the 28S rRNA.
CC FT SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC FT TYPE 1 RIP SUBFAMILY.
CC FT
CC FT This SWISS-PROT entry is copyright. It is produced through a collaboration
CC FT between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC FT the European Bioinformatics Institute. There are no restrictions on its
CC FT use by non-profit institutions as long as its content is in no way
CC FT modified and this statement is not removed. Usage by and for commercial
CC FT entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC FT or send an email to license@isb-sib.ch).
CC FT
CC FT EMBL; I34238; -; NOT_ANNOTATED_CDS.
CC FT HSSP; P09989; 1MRJ.

```

```

FT STRAND 260 260
FT STRAND 263 263
FT TURN 266 268
SQ SEQUENCE 289 AA; 31676 MW; 5CE09BB630575BB9 CRC64;
Query Match 33.4%; Score 342; DB 1; Length 289;
Best Local Similarity 39.5%; Pred. No. 1.4e-24;
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;
QY 10 INFTTAGATVOSYTNFIRAVRGRLTTGADVRHEIPVLPNVRVGLPINORFILVELSNHAEI 69
DB 25 VSRFLSGATSSSYGVFISNLKALPNERKL-YDIPLL--RSLPGSQRYALHILHTNAYADE 81
QY 70 SVTLALDVTNATVGVYRAGNSAYFFHPDQEDA-FAITHLFTDVQNRVTFAGGNYDRLE 128
DB 82 TTSVAIDVTNVYINGRAGDTSYFF---NEASATEAAKYVFKDAMRKVTLPLPSGVNRIQ 138
QY 129 OLAGNLRNIELNGPGLBEAISALYYSTGGTQLPTLARSFICICIMISEAARPOVIEGE 188
DB 139 TAAGKIRENIPGLPALDSAITLIFYNAN-----SAASALMVLQISTSEAAARYKFIEQ 193
QY 189 METRI 193
DB 194 ICKRV 198
RESULT 6
RIP2_BRYDI STANDARD; PRT; 282 AA.
AC P98184; Q988J0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein bryodin II precursor (rRNA N-
DE glycosidase) (EC 3.2.2.22) (BD2).
OS Bryonia dioica (Red bryony).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.
OC NCBI_TaxID=3652;
RN [1]
RP SEQUENCE FROM N.A.
RA Siegal C.B., Gawlak S.L., Marquardt H.;
RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant
RL Bryonia dioica.";
RN [2]
RP Patent number US5597569, 28-JAN-1997.
RP SEQUENCE OF 22-42.
RC TISSUE=Root;
RX MEDLINE=95151812; PubMed=7849072;
RA Siegal C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
RA Marquardt H.;
RT "Characterization of ribosome-inactivating proteins isolated from
RT Bryonia dioica and their utility as carcinoma-reactive
RT immunoconjugates.";
RL Bloconj. Chem. 5:423-429 (1994).
CC -!- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
CC -!- PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; I34238; -; NOT_ANNOTATED_CDS.
CC HSSP; P09989; 1MRJ.

```


CC		EMBL; AB006666; BAA21786.1; -	
DR	PIR	JCS606; JCS606.	
DR	PIR	JU0393; JU0393.	
DR	HSSP	P09989; LMRJ.	
DR	InterPro	IPR001574; RIP.	
DR	Pfam	PF00161; RIP. 1	
DR	PRINTS	PR00396; SHIGARICIN	
DR	PROSITE	PS00275; SHIGA_RICIN. 1.	
KW	Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase; Toxin; Signal.		
FT	SIGNAL	1 21 POTENTIAL.	
FT	CHAIN	22 270 KARASURIN-C.	
FT	CHAIN	24 270 KARASURIN-A.	
FT	PROPEP	271 289 REMOVED IN MATURE FORM.	
FT	ACT_SITE	183 BY SIMILARITY.	
SQ	SEQUENCE	289 AA; 31704 MW; 883D3E3242887B26 CRC64;	
	Query Match	33.0%; Score 338; DB 1; Length 289;	
	Best Local Similarity	40.0%; Pred. No. 3.3e-24;	
	Matches	74; Conservative 46; Mismatches 53; Indels 12; Gaps 5	
Qy	10	INTTTAGATVOSTNFTFRAVRGRLLTGGADYRHETPLPNRVGLGFLINORFIIVLSNHAEI 69	
Dd	25	VSFRLSGATSSYGVFISNRKALPYERKL-YDIPELL--RSTLPQSQRYSALHILTNYADE 81	
Qy	70	SVTLALDVNTAYVVGVEAGNSAYFFHPDNQEDA-EAITHLFTDVQNRYTFAFGGNYDRLE 128	
Dd	82	TTSVALDVNTVVMGYRAGTSTVFF---NEASATEAAKYVKDAKKVTLPYSGNYERLQ 138	
Qy	129	OLAGNLRENIELNGPLUEAISALYYSTGTQTPTLARSLFIICIMISEAARFOYIRGE 188	
Dd	139	IAAGKIRENIPLGLPALDSAITLTLYNNAN----SAASALMWLIQTSEAAKYKFIEQQ 193	
Qy	189	MRTRI 193	
Dd	194	IGRRV 198	
RESULT 8			
ID	NIGB_SAMN1	STANDARD;	PRT; 563 AA.
AC	P33183; P33184; P93542;		
DT	01-OCT-1993 (Rel. 27, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DE	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Nigrin b precursor (Agglutinin V) (SNAP) [Contains: Nigrin b A chain		
DE	(rRNA N-glycosidase) (EC 3.2.2.22); Nigrin b B chain].		
OS	Sambucus nigra (European elder)		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		
OC	Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.		
OX	NCBI_TaxID=4202;		
ON	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Bark;		
RX	MEDLINE=36215449; PubMed=8647092;		
RA	Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;		
RA	"Characterization and molecular cloning of Sambucus nigra agglutinin V		
RT	(nigrin b), a GalNAc-specific type-2 ribosome-inactivating protein		
RT	from the bark of elderberry (Sambucus nigra).";		
RL	Eur. J. Biochem. 237:505-513 (1996).		
RN	[2]		
RP	SEQUENCE OF 26-49 AND 298-321.		
RC	TISSUE=Bark;		
RX	MEDLINE=94003077; PubMed=8400135;		
RA	Girbes T., Citoras L., Ferreras J.M., Rojo M.A., Iglesias R.,		
RA	Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;		
RT	"Isolation and partial characterization of nigrin b, a non-toxic		
RT	novel type 2 ribosome-inactivating protein from the bark of Sambucus		
RT	nigra L.";		
RL	Plant Mol. Biol. 22:1181-1186(1993).		
CC	-!	FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN	

CC PROTEIN SYNTHESIS BUT DOES NOT AFFECT PLANT NOR BACTERIAL PROTEIN
 CC SYNTHESIS. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE
 CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE
 CC BINDING OF NIGRIN B TO THE CELL MEMBRANE THAT PRECEDES
 CC ENDOCYTOSIS.
 CC
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U41299; AAB39475.1; -
 CC
 CC PIR; S37382; S37382; Ricin_B_lectin.
 CC InterPro; IPR000772; Ricin_B_lectin.
 CC InterPro; IPR001574; RIP.
 CC Pfam; PF00652; Ricin_B_lectin; 6.
 CC Pfam; PF00161; RIP; 1.
 CC PRINTS; PR00396; SHIGARICIN.
 CC SMART; SM00458; RICIN; 2.
 CC PROSITE; PS00275; SHIGA_RICIN; 1.
 CC PROSITE; PS0231; RICIN_B_LECTIN; 2.
 CC Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 CC Glycoprotein; Lectin; Signal.
 CC SIGNAL 1 25
 CC CHAIN 26 297 NIGRIN B A CHAIN.
 CC CHAIN 298 563 NIGRIN B B CHAIN.
 CC DOMAIN 305 431 RICIN B-TYPE LECTIN 1.
 CC DOMAIN 434 559 RICIN B-TYPE LECTIN 2.
 CC REPEAT 316 356 1-ALPHA.
 CC REPEAT 357 397 1-BETA.
 CC REPEAT 400 432 1-GAMMA.
 CC REPEAT 445 482 2-ALPHA.
 CC REPEAT 486 524 2-BETA.
 CC REPEAT 527 554 2-GAMMA.
 CC ACT_SITE 188 188 BY SIMILARITY.
 CC DISULFID 274 302 INTERCHAIN (BY SIMILARITY).
 CC DISULFID 319 338 BY SIMILARITY.
 CC DISULFID 360 377 BY SIMILARITY.
 CC DISULFID 448 463 BY SIMILARITY.
 CC DISULFID 489 506 BY SIMILARITY.
 CC CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CONFLICT 39 39 X -> V (IN REF. 2).
 CC SEQUENCE 563 AA; 62300 MW; F250CBE24621BF14 CRC64;
 CC
 CC Query Match 32.2%; Score 330.5; DB 1; Length 563;
 CC Best Local Similarity 39.5%; Pred. No. 3,7e-23;
 CC Matches 75; Conservative 39; Mismatches 61; Indels 15; Gaps 4;
 CC
 CC Qy 7 YPIINFAGTQSYNYFTRAVGRLLTGTADVREHPIVLPNVRGLPINOQFILVELSNH 66
 CC Db 28 YPSVFNLDGAKSATYRDFLSNLRKKTATGTYEYVGLPVLRRSEVQKSRFVLVPLTNY 87
 CC Qy 67 AELSVTLALDVNAYVGYRAGNSAYFPHDPNQDAEAL--THLFTDQVRYFAFGNY 124
 CC Db 88 NGNTVTLAVDVNLYVAFSGNANSYFF-----KDAVEVQKSNLDFVGTQKN-TLSFTGNY 141
 CC Qy 125 DRLEQLAGNLRNLELGNPLLEEAIISALXYVSTGCTQLPTLARSFIIICMISEARFQY 184

Db 142 DNLETAANTRRRESIELGSPDLGATSLYHGD-----SVARSLIVIQMSEARFQY 194
 Qy 185 IEGEMTRIR 194
 Db 195 ISEQEVRSLQ 204
 RESULT 9
 ID ABRA ABRPR STANDARD; PRT; 528 AA.
 AC P11140; P28589;
 DT 01-JUL-1989 (Rel. 11, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Abrin-a precursor [Contains: Abrin-a A chain (RNA N-glycosidase)
 (EC 3.2.2.22); Abrin-a B chain].
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid 1; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 OC NCBI_TaxID=3816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93132798; PubMed=8421313;
 RA Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
 RT "Primary structure of three distinct isoabrinins determined by cDNA
 RT sequencing. Conservation and significance.";
 RL J. Mol. Biol. 229:263-267(1993).
 RN [2]
 RP SEQUENCE OF 1-251.
 RC TISSUE=Seed;
 RA Funatsu G., Taguchi Y., Kamenosono M., Yanaka M.;
 RT "The complete amino acid sequence of the A-chain of abrin-a, a toxic
 RT protein from the seeds of Abrus precatorius.";
 RL Agric. Biol. Chem. 52:1095-1097(1988).
 RN [3]
 RP SEQUENCE OF 1-251 FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=91201329; PubMed=2016300;
 RA Evensen G., Mathiesen A., Sundan A.;
 RT "Direct molecular cloning and expression of two distinct abrin
 RT A-chains.";
 RL J. Biol. Chem. 266:6848-6852(1991).
 RN [4]
 RP SEQUENCE OF 262-528.
 RX MEDLINE=92371656; PubMed=1505674;
 RA Chen Y.-L., Chow L.-P., Tsugita A., Lin J.-Y.;
 RT "The complete primary structure of abrin-a B chain.";
 RL FEBS Lett. 309:115-118(1992).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.14 ANGSTROMS).
 RX MEDLINE=95333188; PubMed=7608980;
 RA Tahirov T.H., Lu T.-H., Liaw Y.-C., Chen Y.-L., Lin J.-Y.;
 RT "Crystal structure of abrin-a at 2.14 A.";
 RL J. Mol. Biol. 250:354-367(1995).
 CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
 CC ABRIN-A IS MORE TOXIC THAN RICIN.
 CC -!- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
 CC PRECEDES ENDOCYTOSIS.
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M98344; AAA32624.1; ALT INIT.
 DR EMBL; X54872; -; NOT_ANNOTATED_CDS.
 DR PIR; S32429; TZLSA.
 DR PDB; 1ABR; 07-FEB-95.
 DR InterPro; IPR000772; Ricin_B_lectin.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00652; Ricin_B_lectin; 6.
 DR Pfam; PF00161; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS0231; RICIN_B_LLECTIN; 2.
 DR PROSITE; PS0275; SHIGA_RICIN; 1.
 KW Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 KW Glycoprotein; Lectin; 3D-structure; Pyrrolidone carboxylic acid.
 FT CHAIN 1 251 ABRIN-A A CHAIN.
 FT PEPTIDE 252 261 LINKER PEPTIDE.
 FT CHAIN 262 528 ABRIN-A B CHAIN.
 FT DOMAIN 273 400 RICIN B-TYPE LECTIN 1.
 FT DOMAIN 403 527 RICIN B-TYPE LECTIN 2.
 FT REPEAT 283 325 1-ALPHA.
 FT REPEAT 326 366 1-BETA.
 FT REPEAT 369 401 1-GAMMA.
 FT REPEAT 414 449 2-ALPHA.
 FT REPEAT 453 492 2-BETA.
 FT REPEAT 495 528 2-GAMMA.
 FT ACT SITE 164 164 BY SIMILARITY.
 FT DISULFID 247 269 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 286 305 BY SIMILARITY.
 FT DISULFID 329 346 BY SIMILARITY.
 FT DISULFID 417 430 BY SIMILARITY.
 FT DISULFID 456 473 BY SIMILARITY.
 FT MOD RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .).
 FT CONFLICT 202 202 MISSING (IN REF. 2).
 FT CONFLICT 298 298 N -> Y (IN REF. 4).
 FT CONFLICT 427 427 M -> L (IN REF. 4).
 FT CONFLICT 457 457 T -> P (IN REF. 4).
 FT CONFLICT 483 483 V -> L (IN REF. 4).
 FT STRAND 5 8
 FT TURN 10 11
 FT HELIX 14 28
 FT STRAND 32 33
 FT TURN 34 35
 FT STRAND 36 38
 FT TURN 42 43
 FT HELIX 47 49
 FT STRAND 51 57
 FT TURN 63 69
 FT TURN 70 72
 FT STRAND 75 79
 FT STRAND 83 86
 FT TURN 88 89
 FT TURN 92 93
 FT HELIX 94 97
 FT TURN 100 101
 FT STRAND 103 106
 FT TURN 113 114
 FT HELIX 115 119
 FT TURN 124 126
 FT STRAND 129 129
 FT HELIX 131 142
 FT TURN 143 144
 FT HELIX 148 167
 FT STRAND 168 168
 FT HELIX 169 180
 FT TURN 181 182

FT STRAND 185 185
 FT HELIX 189 196
 FT TURN 197 197
 FT HELIX 198 207
 FT STRAND 212 220
 FT TURN 222 223
 FT STRAND 226 231
 FT TURN 232 233
 FT HELIX 235 239
 FT STRAND 240 240
 FT STRAND 243 243
 FT STRAND 248 248
 FT STRAND 268 268
 FT STRAND 276 277
 FT STRAND 279 280
 FT HELIX 282 284
 FT STRAND 286 289
 FT HELIX 290 292
 FT TURN 296 297
 FT STRAND 299 303
 FT HELIX 311 313
 FT STRAND 315 317
 FT TURN 319 320
 FT STRAND 322 325
 FT TURN 326 327
 FT STRAND 328 332
 FT TURN 337 338
 FT STRAND 340 344
 FT TURN 346 348
 FT HELIX 351 353
 FT STRAND 355 355
 FT STRAND 357 358
 FT TURN 360 361
 FT STRAND 364 366
 FT TURN 371 374
 FT STRAND 371 374
 FT TURN 380 381
 FT STRAND 383 383
 FT STRAND 385 387
 FT HELIX 393 395
 FT STRAND 398 399
 FT STRAND 406 408
 FT STRAND 410 411
 FT HELIX 413 415
 FT STRAND 417 421
 FT TURN 422 423
 FT STRAND 424 428
 FT TURN 432 433
 FT HELIX 435 437
 FT STRAND 439 441
 FT TURN 443 444
 FT STRAND 447 449
 FT TURN 450 451
 FT STRAND 452 459
 FT TURN 464 465
 FT STRAND 467 472
 FT TURN 474 475
 FT HELIX 478 480
 FT STRAND 483 484
 FT TURN 486 487
 FT STRAND 490 492
 FT TURN 493 496
 FT STRAND 497 501

Query Match 32.1%; Score 329.5; DB 1; Length 528;
 Best Local Similarity 42.8%; Pred. No. 4.2e-23;

Matches 80; Conservative 28; Mismatches 66; Indels 13; Gaps 5;

QY 10 INFTTAGATVQSYTNFIRAVRGRLTTGADVREHPVLPNVRVGLPINQRFILVELSNHAEL 69

Db 5 IKFTEGATQSQYKQFIEALRELRGG--LIHIDPVLDPFTLQERNRITVELSNSDTE 62

QY 70 SVTLALDVTNAYVVGVRAGNSAYFFH--PDNQEDAEAIHTLFTDVQNRVYTFAGGNYDL 127


```

MLA VISAL
ID _MLA VISAL STANDARD; PRT; 254 AA.
AC P81446;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Beta-galactoside specific lectin I A chain (MLA) (ML-I A) (rRNA N-
glycosidase) (EC 3.2.2.22).
OS Viscum album (European mistletoe).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Santalales; Santalaceae; Viscum.
OX NCBI_TaxID=3972;
RN [1]
RP SEQUENCE.
RC STRAIN=Subsp. album;
RX MEDLINE=97134581; PubMed=8980141;
RA Soler M.H.; Stoeva S.; Schwaborn C.; Wilhelm S.; Stiefel T.;
RA Voelter W.;
RT "Complete amino acid sequence of the A chain of mistletoe lectin I.";
RL FEBS Lett. 399:153-157(1996).
CC -1- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S rRNA. THE
B CHAIN BINDS TO CELL RECEPTORS AND PROBABLY FACILITATES THE ENTRY
INTO THE CELL OF THE A CHAIN; B CHAINS ARE ALSO RESPONSIBLE FOR
CELL AGGLUTINATION (LECTIN ACTIVITY).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
specific adenosine on the 28S rRNA.
CC -1- SUBUNIT: Disulfide-linked dimer of A and B chains.
CC -1- PHARMACOLOGICAL: Due to its immunomodulatory effects it is being
studied in clinical trials in cancer patients as it may slow the
growth of cancer cells and be an effective treatment for solid
tumors.
CC -1- MISCELLANEOUS: TWO ISOFORMS OF MLA EXIST: GLYCOSYLATED FORM MLA
AND NON-GLYCOSYLATED FORM MLA'.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 2 RIP SUBFAMILY.
DR PIR; PD0018; PD0018.
DR HSSP; P11140; IABR.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; FALSE NEG.
KW Plant defense; Hydrolase; Toxin; Repeat; Glycoprotein; Lectin.
FT ACT_SITE 165 165
FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .).
FT VARIANT 15 15 E -> D (IN MLA').
FT VARIANT 66 66 V -> I (IN MLA').
FT VARIANT 112 112 N -> T (IN MLA').
FT VARIANT 116 116 P -> T (IN MLA').
FT VARIANT 133 134 DQ -> EE (IN MLA').
FT VARIANT 140 140 T -> S (IN MLA').
FT VARIANT 144 144 F -> Y (IN MLA').
FT VARIANT 151 151 T -> A (IN MLA').
FT VARIANT 179 179 Y -> D (IN MLA').
FT VARIANT 184 184 A -> E (IN MLA').
FT VARIANT 190 190 V -> M (IN MLA').
FT VARIANT 218 218 I -> F (IN MLA').
FT VARIANT 223 224 PP -> ST (IN MLA').
FT VARIANT 231 231 T -> S (IN MLA').
FT VARIANT 235 235 D -> S (IN MLA').
SQ SEQUENCE 254 AA; 28478 MW; 53BAF98D3E0FFE67 CRC64;

Query Match 27.9%; Score 286; DB 1; Length 254;
Best Local Similarity 38.4%; Pred. No. 1.9e-19;
Matches 81; Conservative 36; Mismatches 56; Indels 38; Gaps 10;

QY 14 TAGATVQSYTNFIRAVRGRLTTCADVRHEIPVLPNRVGLPIN--QRFTLVLSNHAELSV 71
DB 9 THQTGEEYFRFTLLRDYVSSGS--FSNEIFLL-RQSTIVSDAQRFVLVLTGQGVSV 66
QY 72 TLALDVTNAYVVGVRAGNSAYFFHPDNCQDAEAITHLFTDQVNRVYTFAGGNYDRLEQLA 131

RESULT 14
RIPA LUFUCY
ID RIPA LUFUCY STANDARD; PRT; 277 AA.
AC Q00465;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein luffin-alpha precursor (rRNA
N-glycosidase) (EC 3.2.2.22).
DE Luffa cylindrica (Smooth loofah) (Sponge gourd).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Luffa.
OX NCBI_TaxID=3670;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=92288316; PubMed=1600156;
RA Kataoka J.; Habuka N.; Miyano M.; Masuta C.; Koizumi A.;
RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
inactivating protein from Luffa cylindrica.";
RL Plant Mol. Biol. 18:1199-1202(1992).
CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
specific adenosine on the 28S rRNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; X62371; CAA44229.1; -.
CC PIR; S22494; S22494.
CC HSSP; P16094; LAHC.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00161; RIP; 1.
CC PRINTS; PR00396; SHIGARICIN.
CC PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
Toxin; Signal.
FT SIGNAL 1 19 RIBOSOME-INACTIVATING PROTEIN LUFFIN-
FT CHAIN 20 277 ALPHA.
FT ACT_SITE 179 179 BY SIMILARITY.
SQ SEQUENCE 277 AA; 30212 MW; EA17C27998C25AC CRC64;

Query Match 27.8%; Score 284.5; DB 1; Length 277;
Best Local Similarity 33.5%; Pred. No. 2.9e-19;
Matches 63; Conservative 45; Mismatches 69; Indels 11; Gaps 3;

QY 10 INTTACATVQSYTNFIRAVRGRLTTCADVRHEIPVLPNRVGLPINQRFTLVLSNHAEL 69
DB 22 VRFSLGSSSTSYSKFGIDRLKALPSNGTVYNTLLISSASGA---SRYTILMTLSNDGK 78
QY 70 SVTLALDVTNAYVVGVRAGNSAYFFHPDNCQDAEAITHLFTDQVNRVYTFAGGNYDRLEQ 129
DB 79 AITVADVNTVNYVVGVRAGNSAYFFHPDNCQDAEAITHLFTDQVNRVYTFAGGNYDRLEQ 135

```

QY 130 LAGNLRNIEGNGPLEEARISALYYSTGGTQLTPLARSFICICMISEARFQYIEGEM 189
Db 136 AAGKIREKIPGLFPALDSAITTLFHYDS-----TAAAAAFVITQTTAEASRFKYIEGQI 190
QY 190 RTRIRYNR 197
Db 191 IERISKNO 198

RESULT 15

RIP2 MOWBA
ID_RIP2_MOWBA STANDARD; PRT; 286 AA.
AC P29339;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ribosome-inactivating protein momordin II precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Momordica balsamina (Bitter melon) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3672;
FN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=Seed;
RX MEDLINE=93027170; PubMed=1408771;
RT Ortiga M., Better M.;
RA "Momordin II, a ribosome inactivating protein from Momordica
RT balsamina, is homologous to other plant proteins.";
RL Nucleic Acids Res. 20:4662-4662(1992).
CC -! CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -! SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z12175; CAA78166.1; -.
DR PIR; S25560; S25560.
DR PDB; 1CF5; 07-JUN-99.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;
FT 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 286 RIBOSOME-INACTIVATING PROTEIN MOMORDIN
FT ACT SITE 181 181 BY SIMILARITY.
FT SEQUENCE 286 AA; 32031 MW; 3B89FF1AE6B25986 CRC64;

Query Match 27.2%; Score 279; DB 1; Length 286;
Best Local Similarity 35.3%; Pred. No. 9.9e-19;
Matches 66; Conservative 42; Mismatches 67; Indels 12; Gaps 5;
QY 10 INFTTAGATVQSYTNFIRAVRGRLTGTADVRHEIPVLPNVRGLPINORFILVELSNHAEI 69
Db 25 VNFDLSTAKTYTKFIEDFRATLPESHKV-YDIPLLYSTIS--DSRRFILLDLTSYAYE 81
QY 70 SYTLALDVNAYVVGVRAGNSAYFFHPDQDEAEATHLFTDVQNRYPFAFCGNYDRLEQ 129
Db 82 TISVAIDVTNVVWVAYRTRDVSIFYF---KESPEAYNILLFKGTR-KITLPYTGNYENLQT 137
QY 130 LAGNLRNIEGNGPLEEARISALYYSTGGTQLTPLARSFICICMISEARFQYIEGEM 189

Db 138 AAHKIRENIDGLPALSSAITTLFYNA-----QSAPSALLVLIQTAEARFYIERHV 192
QY 190 RTRIRYN 196
Db 193 AKYVATN 199

Search completed: February 10, 2004, 16:23:23
Job time : 7.63825 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:00 ; Search time 26.553 Seconds
(without alignments)
1933.961 Million cell updates/sec

Title: US-10-083-336A-5

Perfect score: 1025

Sequence: 1 MIFPKQYPIINFITAGATVQ.....ARFQVIEGEMTRIRYNRRS 199

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23.*

- 1: sp archea.*
- 2: sp bacteria.*
- 3: sp fungi.*
- 4: sp human.*
- 5: sp invertebrate.*
- 6: sp mammal.*
- 7: sp mhc.*
- 8: sp organelle.*
- 9: sp phage.*
- 10: sp plant.*
- 11: sp rodent.*
- 12: sp virus.*
- 13: sp vertebrate.*
- 14: sp unclassified.*
- 15: sp rvirus.*
- 16: sp bacteriaph.*
- 17: sp archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1017	99.2	541	Q41174	ricinus com
2	401.5	39.2	580	Q94BW3	Q94bw3 cinnamomum
3	397.5	38.8	580	Q94BW4	Q94bw4 cinnamomum
4	397.5	38.8	581	Q94BW5	Q94bw5 cinnamomum
5	395.5	38.6	549	Q94V22	Q94v22 cinnamomum
6	350.5	34.2	563	Q04367	Q04367 sambucus ni
7	347.5	33.9	564	Q94VR2	Q94vr2 sambucus eb
8	344.5	33.6	528	Q06076	Q06076 abrus preca
9	340	33.2	289	Q04KE4	Q04ke4 trichosan
10	338	33.0	247	Q04LRE3	Q04lre3 trichosan
11	338	33.0	289	Q41216	Q41216 trichosan
12	336.5	32.8	252	Q038760	Q038760 abrus preca
13	330.5	32.2	563	Q045S2	Q045s2 sambucus ni
14	330.5	32.2	563	Q08GT32	Q08gt32 sambucus ni
15	329.5	32.1	252	Q038761	Q038761 abrus preca
16	324	31.6	270	Q08LPV7	Q08lpv7 trichosan

ALIGNMENTS

RESULT 1

Q41174 ID Q41174 PRELIMINARY; PRT; 541 AA.

AC Q41174; DT 01-NOV-1996 (TREMREL. 01, Created)
DT 01-NOV-1996 (TREMREL. 01, Last sequence update)
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE Proricin A chain (EC 3.2.2.22) (rRNA N-glycosidase)
DE (Fragment).
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosida I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92338377; PubMed=1633311;
RA Roberts L.M., Tregear J.W., Lord J.M.;
RT "Molecular cloning of ricin."
RL Targeted Diagn. Ther. 7:81-97(1992).
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC EMBL; S40366; AAB22582.1; -.
DR HSSP; P02879; 1BR6.
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR InterPro; IPR001400; Somatotropin.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS00231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
DR KX Hydrolase; Toxin.
FT NON TER 1
SQ SEQUENCE 541 AA; 60281 MW; 2B7B2CDEF1F2E9D9 CRC64;

Q96236 abrus preca
Q96237 abrus preca
Q04071 sambucus ni
Q41611 trichosan
Q96235 abrus preca
Q9669 abrus preca
Q04072 sambucus ni
Q00980 luffa cylin
Q41358 sambucus ni
Q81KQ5 viscum albu
Q22415 sambucus ni
Q9M634 polygonatum
Q81KQ4 viscum albu
Q81KQ6 viscum albu
Q9FSH2 momordica c
Q8452 jatropa cu
Q8W243 viscum albu
Q9FUV7 momordica c
P93543 sambucus ni
Q8RXH6 viscum albu
Q8RXH7 viscum albu
Q41257 momordica c
Q04358 iris hollan
Q8W2e7 iris hollan
Q04356 iris hollan
Q9M653 polygonatum
Q8GZN9 euphorbia s
Q8VYU0 jatropa cu
Q8W2e8 iris hollan

```
Query Match          99.2%; Score 1017; DB 10; Length 541;
Best Local Similarity 99.5%; Pred. No. 1.9e-87;
Matches 197; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILV 61
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILV 60

Qy 62 ELSNHAELSVTLADVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFG 121
Db 61 ELSNHAELSVTLADVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFG 120

Qy 122 GNYDRLEQLAGNRENIELGNPLEEALSALYYSTGQTLPRLARSFIIQIMISEAAR 181
Db 121 GNYDRLEQLAGNRENIELGNPLEEALSALYYSTGQTLPRLARSFIIQIMISEAAR 180

Qy 182 FQYIEGEMRTIRYNRRS 199
Db 181 FQYIEGEMRTIRYNRRS 198

RESULT 2
Q94BW3 PRELIMINARY; PRT; 580 AA.
AC Q94BW3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin III precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC EMBL; AY039803; AAK82460.1; -
CC InterPro; IPR000772; Ricin_B_lectin.
CC DrPro; IPR001574; RIP.
CC Pfam; PF00652; Ricin_B_lectin; 6.
CC PRINTS; PR00396; SHIGARICIN.
CC SMART; SM00458; RICIN; 2.
CC PROSITE; PS50231; RICIN_B_LECTIN; 2.
CC Hydroxylase; Signal; Toxin.
KW SIGNAL 1 32
FT CHAIN 33 580
FT SIGNAL 1 32
FT CHAIN 33 580
SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7FB558 CRC64;

Query Match          39.2%; Score 401.5; DB 10; Length 580;
Best Local Similarity 50.0%; Pred. No. 2.8e-29;
Matches 94; Conservative 28; Mismatches 59; Indels 7; Gaps 5;

Qy 7 YPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVLSN- 65
Db 33 YQTVTFITKNATKTSYTOFIEALRAQLASGEE-PHGIPVMRDSSTVPDSKRFILVLSNW 91

Qy 66 HAELSVTALDVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 125
Db 92 AADSPVTLAVDVNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 148

Qy 126 RLEQLAGNRENIELGNPLEEALSALYYSTGQTLPRLARSFIIQIMISEAARFYI 185
Db 149 DLERVAGERREILLGMDPLENALSAL--NQQRALARSLLIVVIQWVAEAVRFRFI 206

Query Match          39.2%; Score 401.5; DB 10; Length 580;
Best Local Similarity 50.0%; Pred. No. 2.8e-29;
Matches 94; Conservative 28; Mismatches 59; Indels 7; Gaps 5;

Qy 7 YPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVLSN- 65
Db 33 YQTVTFITKNATKTSYTOFIEALRAQLASGEE-PHGIPVMRDSSTVPDSKRFILVLSNW 91

Qy 66 HAELSVTALDVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 125
Db 92 AADSPVTLAVDVNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 148

Qy 126 RLEQLAGNRENIELGNPLEEALSALYYSTGQTLPRLARSFIIQIMISEAARFYI 185
Db 149 DLERVAGERREILLGMDPLENALSAL--NQQRALARSLLIVVIQWVAEAVRFRFI 206
```

```
Qy 186 EGEWRTRI 193
Db 207 EYRVRESI 214

RESULT 3
Q94BW4 PRELIMINARY; PRT; 580 AA.
AC Q94BW4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin II precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC EMBL; AY039802; AAK82459.1; -
CC InterPro; IPR000772; Ricin_B_lectin.
CC DrPro; IPR001574; RIP.
CC Pfam; PF00652; Ricin_B_lectin; 6.
CC PRINTS; PR00396; SHIGARICIN.
CC SMART; SM00458; RICIN; 2.
CC PROSITE; PS50231; RICIN_B_LECTIN; 2.
CC Hydroxylase; Signal; Toxin.
KW SIGNAL 1 32
FT CHAIN 33 580
FT SIGNAL 1 32
FT CHAIN 33 580
SQ SEQUENCE 580 AA; 64265 MW; 37E4289ECCE0CBFF CRC64;

Query Match          38.8%; Score 397.5; DB 10; Length 580;
Best Local Similarity 49.5%; Pred. No. 6.7e-29;
Matches 93; Conservative 30; Mismatches 58; Indels 7; Gaps 5;

Qy 7 YPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVLSN- 65
Db 33 YQTVTFITKNATKTSYTOFIEALRAQLASGEE-PHGIPVMRDSSTVPDSKRFILVLSNW 91

Qy 66 HAELSVTALDVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 125
Db 92 AADSPVTLAVDVNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 148

Qy 126 RLEQLAGNRENIELGNPLEEALSALYYSTGQTLPRLARSFIIQIMISEAARFYI 185
Db 149 DLERVAGERREILLGMDPLENALSAL--WTSNINQQRALARSLLIVVIQWVAEAVRFRFI 206

Query Match          38.8%; Score 397.5; DB 10; Length 580;
Best Local Similarity 49.5%; Pred. No. 6.7e-29;
Matches 93; Conservative 30; Mismatches 58; Indels 7; Gaps 5;

Qy 7 YPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVLSN- 65
Db 33 YQTVTFITKNATKTSYTOFIEALRAQLASGEE-PHGIPVMRDSSTVPDSKRFILVLSNW 91

Qy 66 HAELSVTALDVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 125
Db 92 AADSPVTLAVDVNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 148

Qy 126 RLEQLAGNRENIELGNPLEEALSALYYSTGQTLPRLARSFIIQIMISEAARFYI 185
Db 149 DLERVAGERREILLGMDPLENALSAL--WTSNINQQRALARSLLIVVIQWVAEAVRFRFI 206

Query Match          38.8%; Score 397.5; DB 10; Length 580;
Best Local Similarity 49.5%; Pred. No. 6.7e-29;
Matches 93; Conservative 30; Mismatches 58; Indels 7; Gaps 5;

Qy 7 YPIINFTTAGATVQSYTNFIRAVRGLTTGADVRRHEIPVLPNRVGLPINQRFILVLSN- 65
Db 33 YQTVTFITKNATKTSYTOFIEALRAQLASGEE-PHGIPVMRDSSTVPDSKRFILVLSNW 91

Qy 66 HAELSVTALDVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 125
Db 92 AADSPVTLAVDVNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTAFGNGYD 148

Qy 126 RLEQLAGNRENIELGNPLEEALSALYYSTGQTLPRLARSFIIQIMISEAARFYI 185
Db 149 DLERVAGERREILLGMDPLENALSAL--WTSNINQQRALARSLLIVVIQWVAEAVRFRFI 206
```

DR	InterPro; IPR000772; Ricin_B_lectin.	
DR	InterPro; IPR001574; RIP.	
DR	Pfam; PF00652; Ricin_B_lectin; 5.	
DR	Pfam; PF00161; RIP; I_	
DR	PRINTS; PR00396; SHIGARICIN.	
DR	SMART; SM00458; RICIN; 2.	
DR	PROSITE; PS0231; RICIN_B_LECTIN; 2.	
KW	Hydrolase; Toxin.	
FT	NON TER	1
SQ	SEQUENCE 549 AA; 60648 MW; 02607FE607CA44B0 CRC64;	
Query Match		
Best Local Similarity 50.0%; Score 395.5; DB 10; Length 549;		
Matches 94; Conservative 27; Mismatches 60; Indels 7; Gaps 5;		
QY	7 YPIINFETGATVQSYTNFIRAVRGRLITTGADVRHEIPVLPNRVGLPINORFILVELSN-	65
DB	1 YQVVTITTKATKTSYTFEALRAQLASGE-PhGIPVWRERSTVPDSKRFILVELSNW	59
QY	66 HAEISVTLALDVNNAVGVGRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTTAFGGNYD	125
DB	60 AADSPVTLADVNTNAVAVYRTGQSFFLRDNP--PAIENLLPDTK-RYTFPPFSGSYT	116
QY	126 RLRLQAGNRETELGNGPLEEALSAIYYSTGGTQLPRLARSFILCIOMISAAARQYI	185
DB	117 DLEGVAGEREEILLGMDPLENAISALWISNL--NOORALARSLLIVVIQWVAEAVRPF	174
QY	186 EGEEMRTRI	193
DB	175 EYRVGSI	182
RESULT 6		
004367	PRELIMINARY; PRT: 563 AA.	
ID	004367	
AC	004367;	
DT	01-JUL-1997 (TREMBLrel. 04, Created)	
DT	01-JUL-1997 (TREMBLrel. 04, Last sequence update)	
DT	01-OCT-2002 (TREMBLrel. 22, Last annotation update)	
DE	Ribosome inactivating protein precursor (EC 3.2.2.22) (rRNA	
DE	N-glycosidase).	
OS	Sambucus nigra (European elder).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
OC	Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.	
NCBI	TaxID=4202;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=98112023; PubMed=9450339;	
RA	Van Damme E.J., Roy S., Barre A., Rouge P., Van Leuven F.,	
RA	Peumans W.J.;	
RT	"The major elderberry (Sambucus nigra) fruit protein is a lectin	
RT	derived from a truncated type 2 ribosome-inactivating protein.";	
RL	Plant J. 12:1251-1260(1997)	
CC	-1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE	
CC	SPECIFIC ADENOSINE ON THE 28S RNA.	
CC	-1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.	
DR	EMBL; U76524; AAC15886.1; -.	
DR	HSP; P02879; 2AAL.	
DR	InterPro; IPR000772; Ricin_B_lectin.	
DR	InterPro; IPR001574; RIP.	
DR	Pfam; PF00652; Ricin_B_lectin; 6.	
DR	Pfam; PF00161; RIP; I_	
DR	PRINTS; PR00396; SHIGARICIN.	
DR	SMART; SM00458; RICIN; 2.	
DR	PROSITE; PS0231; RICIN_B_LECTIN; 2.	
DR	PROSITE; PS00275; SHIGA RICIN; 1.	
KW	Hydrolase; Signal; Toxin.	
FT	SIGNAL	1 25
FT	CHAIN	26 297
FT	CHAIN	298 563
SQ	SEQUENCE 563 AA; 62336 MW; 3ED2B6C08E796205 CRC64;	
POTENTIAL.		
RIBOSOME INACTIVATING PROTEIN, A CHAIN.		
RIBOSOME INACTIVATING PROTEIN, B CHAIN.		

QY	125	DRLQLAGNLRENIELGNGPLEEAISALYYVSGTGLPTLARSFTICIMISEARFOY	184
Db	142	DNLTAAAGTRRESIELGPNPLDGAITSLWY--DGG-----VAKSLVLIIQWPEAARFY	194
QY	185	IEGEWRTRIR	194
Db	195	IEQEVRSLQ	204
RESULT 8			
ID	Q06076	PRELIMINARY;	PRT; 528 AA.
ID	Q06076;		
AC	01-NOV-1996 (TREMBLrel. 01, Created)		
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)		
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)		
DE	Abrin-d (EC 3.2.2.22) (r3NA N-glycosidase) (Fragment).		
OS	Abrus precatorius (Indian licorice) (Crab's eye).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eurosid 1; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.		
OX	NGBI_TaxID=3816;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=93132798; PubMed=8421313;		
RA	Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;		
RA	"Primary structure of three distinct isoabrins determined by cDNA		
RT	sequencing: conservation and significance.";		
RL	J. Mol. Biol. 229:263-267(1993)		
CC	-!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE		
CC	SPECIFIC ADENOSINE ON THE 28S RRNA.		
CC	-!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.		
CC	EMBL; N98346; AAA32626.1; -.		
DR	HSP; P11140; IABR.		
DR	InterPro; IPR000772; Ricin_B_lectin.		
DR	InterPro; IER001574; RIP.		
DR	Pfam; PF00652; Ricin_B_lectin; 6.		
DR	Pfam; PF00161; RIP; 1.		
DR	PRINTS; PR00396; SHIGARICIN.		
DR	SMART; SMO0458; RICIN; 2.		
DR	PROSITE; PS0231; Ricin B LECTIN; 2.		
DR	PROSITE; PS00275; SHIGA_RICIN; 1.		
KW	Hydrolase; Toxin.		
FT	NON TER		
FT	NON TER		
FT	NON TER		
SEQ	SEQUENCE 528 AA; 58870 MW; 62ED42FB8FF60F8 CRC64;		
Query Match			
Best Local Similarity 33.6%; Score 344.5; DB 10; Length 528;			
Matches 85; Conservative 24; Mismatches 71; Indels 9; Gaps 4			
QY	6	QYPINETTAGATVQSNTFNIRAVRGRLTTGADVRRHEIPVLPNRVGLEPINORFILVELSN	65
Db	1	QDOVIKKTTEGATSQSQKFTEALRQLTGG--LIHDIPVLPDPPTVEERNRYITVELSN	58
QY	66	HAELSVTLADVTWAYVGVYRAGNSAYFEHFEDNQEDAETHLEDVNQRVTFAPGGND	125
Db	59	SERSEIEVGIDVTWAYVVAYRAGSQSYFL---RDAPASATYLEPGTO-RISLRDGSIG	114
QY	126	RLEQLAGNLRENIELGNGPLEEAFSAIYYVSGTGLPTLARSFTICIMISEARFOY	185
Db	115	DLERWAHQTRREISIGLQALTAHS---FLRSGASNDEKARTLLVIIQMASEAARYRCI	171
QY	186	EGEWRTIR	194
Db	172	SNRVGVSR	180
RESULT 9			
Q94KE4		PRELIMINARY;	PRT; 289 AA.
ID	Q94KE4		
AC	Q94KE4;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		

```

DR Pfam: PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.
FT NON_TER 1 247
FT NON_TER 247 247
SQ SEQUENCE 247 AA; 27199 MW; 89811AC32892F03F CRC64;

Query Match
Best Local Similarity 33.0%; Score 338; DB 10; Length 247;
Matches 74; Conservative 46; Mismatches 53; Indels 12; Gaps 5;

QY 10 INFTTAGATVQSYNTFIRAVRGRLTTGADVRRHEIPVLPRNVGLPINQRFILVELSNHAEL 69
DB 2 VSPFLSGATSSSYGVFISNLRKALPNRKLYDYIPLL--RSTLPGSQRYALVHLTNYADE 58
QY 70 SVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRYYTFAFGNYDRLE 128
DB 59 TISVAIDVTNVYMGYRAGDTSYFF--NEASATEAAKYVFKDKRVTLPYSGNYERLQ 115
QY 129 QLAGNLRNIELGNGPLEEPAISALYYSTGTQLPTLARSFIICIQMISPAARFQYIEGE 189
DB 116 IAAKIRENIPLGLPALDSAITTLFYNNAN-----SAASALMVLIOSTSEAAKYKFIEQQ 170
QY 189 MRTRI 193
DB 171 IGRKV 175

RESULT 11
Q41216 PRELIMINARY; PRT; 289 AA.
ID Q41216
AC Q41216;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Trichosanthin (EC 3.2.2.22) (rRNA N-glycosidase).
GN TRICHOSANTHIN, TCS.
OS Trichosanthin kiliowii (Mongolian snake-gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
OX NCBI_TaxId=3677;
RN [1]
RP SEQUENCE FROM N.A.
RM MEDLINE=94271613; PubMed=8003348;
RT Zheng H., Wang B., Shaw P., Yeung H.;
RT "[Cloning and DNA sequencing of the gene encoding trichosanthin].";
RL I Chuan Hsueh Pao 21:42-51(1994).
CC -1- CATALYTIC ACTIVITY: ENDODYRDLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; S70176; AAB31048.1; -.
DR HSSP; P09889; 1MRJ.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Hydrolase; Toxin.
SQ SEQUENCE 289 AA; 31650 MW; 286AC14D8BCA175 CRC64;

Query Match
Best Local Similarity 39.8%; Score 338; DB 10; Length 289;
Matches 73; Conservative 46; Mismatches 54; Indels 12; Gaps 5;

QY 10 INFTTAGATVQSYNTFIRAVRGRLTTGADVRRHEIPVLPRNVGLPINQRFILVELSNHAEL 69
DB 25 VSPFLSGATSSSYGVFISNLRKALPNRKLYDYIPLL--RSSLPGSQRYALVHLTNYADE 81
QY 70 SVTLALDVTNNAVYVGYRAGNSAYFFHPDNOEDA-EAITHLFTDVQNRYYTFAFGNYDRLE 128
DB 82 TISVAIDVTNVYMGYRAGDTSYFF--NEASATEAAKYVFKDKRVTLPYSGNYERLQ 138


```

QY 129 OLAGNRENTLGNGLPEEAISALYYSTGGTQPLTLARSFIICIMISEARFQIEG 188
 Db 139 TAGKIRENIPGLPALDSAITLIFYNNAN-----SASALMVLIOSTSEARFKIEQ 193

QY 189 MRTRI 193
 Db 194 IGKVV 198

RESULT 12
 Q38760 ID Q38760 PRELIMINARY; PRT; 252 AA.
 AC Q38760;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Abirin-E (RRNA N-glycosidase) (EC 3.2.2.22) (Fragment).
 GN RIP.
 OS Abus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
 OX NCBI_TaxID=3816;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=LEAF;
 RX MEDLINE=91201329; PubMed=2016300;
 RA Evensen G., Mathiesen A., Sundan A.;
 RT "direct molecular cloning and expression of two distinct abrin A-
 RT chains.";
 RL J. Biol. Chem. 266:6848-6852(1991).
 CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
 CC ABRIN-A IS MORE TOXIC THAN RICIN.
 CC -!- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
 CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
 CC PRECEDES ENDOCYTOSIS.
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
 CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -!- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING
 CC PROTEINS. BELONGS TO TYPE 2 RIP.
 DR EMBL; X54872; CRA38654.1; -;
 DR HSPF; P11140; IABR.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00161; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 KW Hydrolase; Glycosidase; Toxin; Repeat; Glycoprotein; Lectin.
 FT CHAIN 1 252 ABRIN E, A CHAIN (BY SIMILARITY).
 FT NON_TER 252 252
 SQ SEQUENCE 252 AA; 28309 MW; BBFC846B9E92B5DE CRC64;

Query Match 32.8%; Score 336.5; DB 10; Length 252;
 Best Local Similarity 44.9%; Pred. No. 1.2e-23;
 Matches 83; Conservative 24; Mismatches 69; Indels 9; Gaps 4;

QY 10 INFTAGTAVSYTNFIRAVRGLTTCADVRHEIPVLPNRVGLPINORFILVELSNHAE 69
 Db 6 IKFSPEGATSSQYKOFIEALRERLGG--LIHIDIPVLRDPTTVERNRYITVELSNRE 63

QY 70 SVTLALDVNAYVGYRAGNSAYFFHPDQEDAEAI--THLFTDVQNRYYTFAFGNVDLEQ 129
 Db 64 SIEVGIDVTNAYVAYRAGNSQSYFL---RDPASASTYLTGTQ-RYSLRFDGSGYGLER 119

QY 130 LAGNRENTLGNGLPEEAISALYYSTGGTQPLTLARSFIICIMISEARFQIEGEM 189
 Db 120 WAHQTRQISLGLQALTHAIS---FLRSGANDEKARTLIVIQMASEARFYISNRV 176

QY 190 RTRIR 194

Db 177 GVSIR 181

RESULT 13
 Q945S2 ID Q945S2 PRELIMINARY; PRT; 563 AA.
 AC Q945S2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA
 DE N-glycosidase).
 GN AVL.
 OS Sambucus nigra (European elder).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.
 OX NCBI_TaxID=4202;
 [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Leaf;
 RA Van Damme E.J.M.;
 RT "Characterization and cloning of lectins and ribosome-inactivating
 RT proteins from Sambucus nigra leaves.";
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 DR EMBL; AF409135; AAL04123.1; -;
 DR InterPro; IPR000772; Ricin_B_lectin.
 DR Pfam; PF00652; Ricin_B_lectin; 6.
 DR PRINTS; PR00161; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 KW Hydrolase; Toxin.
 SQ SEQUENCE 563 AA; 62242 MW; 07F7CBEDCF33BF10 CRC64;

Query Match 32.2%; Score 330.5; DB 10; Length 563;
 Best Local Similarity 39.5%; Pred. No. 1.4e-22;
 Matches 75; Conservative 39; Mismatches 61; Indels 15; Gaps 4;

QY 7 YPIINFTAGTAVSYTNFIRAVRGLTTCADVRHEIPVLPNRVGLPINORFILVELSNH 66
 Db 28 YPSVSNLDCAKSATYRDFLSNLRKTVATGYEVNGLPVLRRSEVQVKSRFVLPLTNY 87

QY 67 AELSVTALDVNAYVGYRAGNSAYFFHPDQEDAEAI--THLFTDVQNRYYTFAFGN 124
 Db 88 NGNTVTILADVNTNLYVVPFSGNANSYFF-----KDTEVQKSNLFVGTQKN-TLSFTGNY 141

QY 125 DRLEQLAGNRENTLGNGLPEEAISALYYSTGGTQPLTLARSFIICIMISEARFQY 184
 Db 142 DNLETAANTRESEITELGSPFLDGAITSYHGD-----SVARSLVVIQMVSEARFRY 194

QY 185 IEGEMRTRIR 194
 Db 195 IEGEVRRSLQ 204

RESULT 14
 Q8GT32 ID Q8GT32 PRELIMINARY; PRT; 563 AA.
 AC Q8GT32;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Type 2 ribosome-inactivating protein nigrin 1 precursor
 DE (EC 3.2.2.22).
 OS Sambucus nigra (European elder).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Search completed: February 10, 2004, 16:26:37
Job time : 29.553 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:20 ; Search time 32.2239 Seconds
(without alignments)
930.966 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971

Sequence: 1 MIFPKQVPIINFTAGATVQ.....ARFQYIEGEMTRVYNRS 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	951	97.9	267	14	AA1980.DAT:*
2	951	97.9	267	16	AA1981.DAT:*
3	951	97.9	290	18	AA1982.DAT:*
4	951	97.9	290	18	AA1983.DAT:*
5	951	97.9	332	8	AA1984.DAT:*
6	951	97.9	332	8	AA1985.DAT:*
7	951	97.9	332	10	AA1986.DAT:*
8	951	97.9	554	16	AA1987.DAT:*
9	951	97.9	562	10	AA1988.DAT:*

10	951	97.9	565	6	AA1989.DAT:*
11	951	97.9	565	22	AA1990.DAT:*
12	951	97.9	565	22	AA1991.DAT:*
13	951	97.9	576	8	AA1992.DAT:*
14	951	97.9	576	18	AA1993.DAT:*
15	951	97.9	576	20	AA1994.DAT:*
16	951	97.9	576	21	AA1995.DAT:*
17	951	97.9	576	22	AA1996.DAT:*
18	951	97.9	576	22	AA1997.DAT:*
19	949	97.7	268	14	AA1998.DAT:*
20	947	97.5	200	9	AA1999.DAT:*
21	946	97.4	534	14	AA2000.DAT:*
22	946	97.4	565	7	AA2001.DAT:*
23	944	97.2	574	8	AA2002.DAT:*
24	942	97.0	574	10	AA2003.DAT:*
25	941	96.9	267	13	AA2004.DAT:*
26	941	96.9	267	21	AA2005.DAT:*
27	940	96.8	332	11	AA2006.DAT:*
28	934	96.2	267	16	AA2007.DAT:*
29	918	94.5	267	14	AA2008.DAT:*
30	861.5	88.7	540	18	AA2009.DAT:*
31	861.5	88.7	540	18	AA2010.DAT:*
32	801.5	82.5	534	8	AA2011.DAT:*
33	718	73.9	280	10	AA2012.DAT:*
34	336	34.6	247	16	AA2013.DAT:*
35	336	34.6	247	21	AA2014.DAT:*
36	336	34.6	248	11	AA2015.DAT:*
37	336	34.6	248	13	AA2016.DAT:*
38	336	34.6	267	18	AA2017.DAT:*
39	336	34.6	267	18	AA2018.DAT:*
40	336	34.6	289	11	AA2019.DAT:*
41	336	34.6	289	13	AA2020.DAT:*
42	336	34.6	289	13	AA2021.DAT:*
43	336	34.6	289	14	AA2022.DAT:*
44	336	34.6	289	15	AA2023.DAT:*
45	336	34.6	289	18	AA2024.DAT:*

ALIGNMENTS

RESULT 1

AA1980.DAT:*

ID AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

AA1980.DAT:*

PT Analogues of type I ribosome inactivating protein - useful as
 PT cytotoxic agents, immuno toxins for treating autoimmune diseases,
 PT cancer, graft versus host disease and selective cell killing in-vivo
 XX
 XX
 XX Claim 1; Page 92; 163pp; English.
 CC The invention covers analogues of Type I RIPs. Ricin is a Type II
 CC RIP whose A chain is homologous to plant type I RIPs. The analogues
 CC of the invention have a cysteine available for intermolecular
 CC disulphide bonding at an amino acid position corresp. to a position
 CC not naturally available for bonding; the cys residue is located in
 CC the C-terminal region of the analogue between a position corresp. to
 CC amino acid 251 and the C-terminus of ricin A chain. The analogues are
 CC pref. joined via a disulphide linkage to a molecule which specifically
 CC binds to a target cell, e.g. an antibody fragment.
 CC (Updated on 09-JAN-2003 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX
 SQ Sequence 267 AA;

Query Match 97.9%; Score 951; DB 14; Length 267;
 Best Local Similarity 94.9%; Pred. No. 1.3e-93;
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps .1;
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
 DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINQRFILV 60
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNRYYTFAFG 111
 DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNRYYTFAFG 120
 QY 112 GNYDRLEQLAGNIRENIELGNGLPEEAISALYYSTGGTQPTLARSFIIICMISEAAAR 171
 DB 121 GNYDRLEQLAGNIRENIELGNGLPEEAISALYYSTGGTQPTLARSFIIICMISEAAAR 180
 QY 172 FOYIEGEMRTRIRYNRS 189
 DB 181 FOYIEGEMRTRIRYNRS 198

RESULT 2
 AAR63902
 ID AAR63902 standard; protein; 267 AA.
 XX
 AC AAR63902;

XX
 XX 25-MAR-2003 (updated)
 DT 27-JUL-1995 (first entry)
 DT
 XX
 XX Ricin A-chain (RTA).
 DE
 XX Ricin A chain; RTA; ribosome-inactivating proteins; RIPs;
 KW cytotoxic therapeutic agents; autoimmune disease; cancer;
 KW graft-versus-host disease.
 KW Ricinus communis.
 OS
 XX WO9426910-A1.
 PN
 XX 24-NOV-1994.
 PD
 XX 12-MAY-1994; 94WO-US05348.
 PF
 XX 12-MAY-1993; 93US-0064691.
 PR
 XX (XOMA) XOMA CORP.
 PA
 XX Better MD, Carroll SS, Studnicka GM, Carroll SF;
 PI
 XX WPI; 1995-006804/01.
 DR
 XX Polynucleotide(s) encoding type I ribosome-inactivating proteins
 PT

PT - which are suitable for use as components of cytotoxic
 PT therapeutic agents.

Example 3; Fig 1; 221pp; English.

XX
 XX AAR63902 is the ricin A chain gene product, it is analogous to the
 CC ribosome-inactivating proteins (RIPs) described in AAR63903-R63911.
 CC RIPs are the key components of cytotoxic therapeutic agents (CTAs),
 CC which include gene fusion products and immunoconjugates. CTAs may
 CC be used to selectively eliminate any cell type to which a RIP
 CC component is targeted, by the specific binding capacity of the
 CC second component of the agent. They can be used in the treatment
 CC of diseases where the elimination of a particular cell type is
 CC desired, such as autoimmune disease, cancer and graft-versus-host
 CC disease.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC
 XX
 XX Sequence 267 AA;

Query Match 97.9%; Score 951; DB 16; Length 267;
 Best Local Similarity 94.9%; Pred. No. 1.3e-93;
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
 DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINQRFILV 60
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNRYYTFAFG 111
 DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFPHDQDEAEATHLFTDVQNRYYTFAFG 120
 QY 112 GNYDRLEQLAGNIRENIELGNGLPEEAISALYYSTGGTQPTLARSFIIICMISEAAAR 171
 DB 121 GNYDRLEQLAGNIRENIELGNGLPEEAISALYYSTGGTQPTLARSFIIICMISEAAAR 180
 QY 172 FOYIEGEMRTRIRYNRS 189
 DB 181 FOYIEGEMRTRIRYNRS 198

RESULT 3
 AAR25136
 ID AAR25136 standard; protein; 290 AA.
 XX
 AC AAR25136;

XX
 XX 25-MAR-2003 (updated)
 DT 02-DEC-1997 (first entry)
 DT
 XX
 XX Ricin A-chain ribosome inhibitory protein inactive precursor.
 DE
 XX Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;
 KW internal linker; Barley translation inhibitor; trichosanthin;
 KW Ricin A-chain; Abrin-A chain; Saporin; SUT-1; Luffin A; MAP;
 KW Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;
 KW therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;
 KW post-translational modification; cancer; neoplasia; HIV; AIDS;
 KW human immunodeficiency virus; acquired immune deficiency syndrome.
 XX
 OS Synthetic.
 XX
 XX US5646026-A.
 FN
 XX 08-JUL-1997.
 PD
 XX 07-JUN-1995; 95US-0485286.
 PF
 XX 09-DEC-1992; 92US-0987927.
 PR 11-JUN-1990; 90US-0535636.
 PR 26-JAN-1995; 95US-0378761.
 PR 07-JUN-1995; 95US-0485286.
 XX
 XX (DOWC) DOWELANCO.

XX Hey TD, Morgan AER, Walsh TA;
 XX WPI; 1997-362934/33.
 XX DNA encoding pro-ribosome inactivating proteins - inactive
 XX precursors of ribosome inactivating proteins; can be expressed in
 XX eukaryotic cells without causing cell death
 XX Claim 4; Column 91-94; 186pp; English.
 XX AAW25136 represents a Ricin A-chain ribosome inhibitory protein (RIP)
 XX which was engineered to contain a selectively removable internal peptide
 XX linker sequence separating the alpha and beta units of the RIP. When
 XX separated the two units regain activity and are capable of inactivating
 XX eukaryotic ribosomes and hence preventing protein production. Many
 XX different RIPs may be produced with an internal linker including
 XX maize RIP, Trichosanthin, Ricin A-chain, Abrin-A A-chain and
 XX Saporin. The RIPs can be used in the construction of therapeutic
 XX toxins targeted to specific cells such as tumour cells via the
 XX attachment of a targeting polypeptide, e.g. a monoclonal antibody.
 XX A further use is in HIV therapy (see US4869903). There is interest
 XX in expressing RIP recombinantly in host eukaryotic cells, because of
 XX the capacity to provide correct post-translational processing. However,
 XX RIPs effectively inhibit protein synthesis in eukaryotic cells resulting
 XX in cell death. Since the inactive RIP proteins are not cytotoxic to
 XX eukaryotic cells, they can be recombinantly expressed in such cells and
 XX then converted to active RIP proteins.
 XX (Updated on 25-MAR-2003 to correct PF field.)
 XX Sequence 290 AA;
 Query Match 97.9%; Score 951; DB 18; Length 290;
 Best Local Similarity 94.9%; Pred. No. 1.4e-93;
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFLIV 51
 Db 25 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLTGGADVVRHEIPVLPNRVGLPINORFLIV 84
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 111
 Db 85 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 144
 QY 172 FQYIEGEMRTRIRYNRRS 189
 Db 205 FQYIEGEMRTRIRYNRRS 222
 RESULT 4
 AAW21699
 ID AAW21699 standard; Protein; 290 AA.
 XX AAW21699;
 XX 25-MAR-2003 (updated)
 DT 26-SEP-1997 (first entry)
 XX Ricin A-chain RIP.
 XX pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;
 XX inactivation; eukaryotic ribosome; alpha fragment; beta fragment;
 XX inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;
 XX rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.
 OS Ricinus communis.
 XX Key
 FH Location/Qualifiers
 FT Region 152..162

FT Hey TD, Morgan AER, Walsh TA;
 FT WPI; 1997-309831/28.
 XX Inactive precursor of maize ribosome-inactivating protein - also
 XX chimeric ribosome-inactivating protein precursors containing
 XX internal linker sequences
 XX Claim 2; Column 91-94; 121pp; English.
 XX The sequences given in AAW21698-710 represent Ribosome Inactivating
 XX Proteins (RIP's), which may be used in the construction of the
 XX proRIP of the invention. The proRIP has a selectively removable,
 XX internal peptide linker. The precursor sequence is incapable of
 XX inactivating eukaryotic ribosomes, but can be converted by removal
 XX of the linker into a protein having alpha and beta fragments and being
 XX capable of inactivating eukaryotic ribosomes. RIPs are potent
 XX inhibitors of eukaryotic protein synthesis. They possess a highly
 XX specific N-glycosidase activity which cleaves the glycosidic bond of
 XX adenine 4324 of rat liver ribosomal 28S RNA. RIP's selectively inhibit
 XX cellular proliferation of cells, e.g. cancer cells and HIV-infected T
 XX cells. The inactive proRIP proteins make it possible to provide protein
 XX synthesis inhibitors with uses in practical and improved ways not before
 XX possible. The RIP can be used to make cytotoxic conjugates.
 XX (Updated on 25-MAR-2003 to correct PF field.)
 XX Sequence 290 AA;
 Query Match 97.9%; Score 951; DB 18; Length 290;
 Best Local Similarity 94.9%; Pred. No. 1.4e-93;
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 QY 2 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFLIV 51
 Db 25 IFPKQYPIINFATTAGATVQSYTNFIRAVRGRLTGGADVVRHEIPVLPNRVGLPINORFLIV 84
 QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 111
 Db 85 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 144
 QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGTQPTLARSFIIQMISEAR 171
 Db 145 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGTQPTLARSFIIQMISEAR 204
 QY 172 FQYIEGEMRTRIRYNRRS 189
 Db 205 FQYIEGEMRTRIRYNRRS 222
 RESULT 5
 AAP70097
 ID AAP70097 standard; protein; 332 AA.
 XX AAP70097;
 XX 09-APR-1991 (first entry)
 DT Ricin A.

/note= "Position of possible insertion of internal
 peptide linker sequence"

```

XX Ricin A; Met-aminopeptidase.
XX Escherichia coli.
XX EP219237-A.
XX 22-APR-1987.
XX 19-SEP-1986; 86EP-0307242.
XX 06-MAY-1986; 86US-0860330.
XX 20-SEP-1985; 85US-0778414.
XX (CETU ) CETUS CORP.
XX Benbassat A, Bauer KA, Chang S, Chang SY;
XX WPI; 1987-110172/16.
XX N-PSDB; AAN70152.
XX N-terminal methionine free proteins prodn. - by using host
XX transformed with vector to express a methionine-amino-peptidase
XX Disclosure; Fig. 4; 20pp; English.
XX Ricin A may be produced in a form which lacks an N-terminal Met
XX using Met-aminopeptidase from E.coli.
XX Sequence 332 AA;
XX Query Match 97.9%; Score 951; DB 8; Length 332;
XX Best Local Similarity 94.9%; Pred. No. 1.7e-93;
XX Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
XX 2 IFFKQYPIINFTTAGATVQSYTNFIRAVGRILT-----VLPNRVGLPINQRFILV 51
XX 36 IFFKQYPIINFTTAGATVQSYTNFIRAVGRILTGDVREHIEPVLNVRVGLPINQRFILV 95
XX 52 ELSNHAELSVTLALDVTNNAVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVYTFAG 111
XX 96 ELSNHAELSVTLALDVTNNAVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVYTFAG 155
XX 112 GNYDRLEQLAGNLRNENIELNGPLEEAISALYYVYSTGGTQPTLARSFFIICQMISEAR 171
XX 156 GNYDRLEQLAGNLRNENIELNGPLEEAISALYYVYSTGGTQPTLARSFFIICQMISEAR 215
XX 172 FOYIEGEMRTRIRYNRRS 189
XX 216 FOYIEGEMRTRIRYNRRS 233
XX RESULT 6
XX AAP70838
XX ID AAP70838 standard; protein; 332 AA.
XX AC AAP70838;
XX 25-MAR-2003 (updated)
XX 18-FEB-1991 (first entry)
XX Sequence of Ricinus communis castor beans ricin toxin (RT or ricin)
XX A protein encoded by pRA123.
XX Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
XX plant toxin.
XX Ricinus communis.
XX Key Location/Qualifiers
XX Region 1..32
XX /note="Leader"
XX Region 33..302

```

```

FT Region /note="A-chain"
FT 315..332
FT /note="B-chain"
XX EP237676-A.
XX 23-SEP-1987.
XX 13-NOV-1986; 86EP-0308877.
XX 07-MAR-1986; 86US-0837583.
XX (CETU ) CETUS CORP.
XX (CHIR ) CHIRON CORP.
XX Piatak M;
XX WPI; 1987-265177/38.
XX N-PSDB; AAN70519.
XX New non-glycosylated ricin precursor and toxin etc. - are prepd.
XX by recombinant DNA procedures with specific isolation steps for
XX purer and soluble prods.
XX Disclosure; Fig 1; 112pp; English.
XX The full-length sequences encoding ricin A (AAN70520), ricin D
XX (AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor
XX form were obtd. using messenger RNA to obtain a cDNA library, and
XX then probing the library to retrieve the desired cDNA inserts. The
XX library was probed using the 35-mer given in AAN70514. Figure 4 (see
XX AAN70520, AAN70521, AAN70522) shows the nucleotide sequences of three
XX plasmids contg. cDNA inserts obtd. by probing a cDNA library for
XX sequences encoding ricin B using the probe in AAN70517. The cDNA
XX inserts can be placed into expression vectors. Site-directed
XX mutagenesis may be used to place an ATG start codon and a HindIII
XX site at the beginning of the mature protein (see AAN70518). The
XX coding sequences of the inserts can be ligated into expression
XX vectors contg. the Phoa promoter-operator and leader sequence
XX (AAN70523) and suitable retroregulators.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX Sequence 332 AA;
XX Query Match 97.9%; Score 951; DB 8; Length 332;
XX Best Local Similarity 94.9%; Pred. No. 1.7e-93;
XX Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
XX 2 IFFKQYPIINFTTAGATVQSYTNFIRAVGRILT-----VLPNRVGLPINQRFILV 51
XX 36 IFFKQYPIINFTTAGATVQSYTNFIRAVGRILTGDVREHIEPVLNVRVGLPINQRFILV 95
XX 52 ELSNHAELSVTLALDVTNNAVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVYTFAG 111
XX 96 ELSNHAELSVTLALDVTNNAVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVYTFAG 155
XX 112 GNYDRLEQLAGNLRNENIELNGPLEEAISALYYVYSTGGTQPTLARSFFIICQMISEAR 171
XX 156 GNYDRLEQLAGNLRNENIELNGPLEEAISALYYVYSTGGTQPTLARSFFIICQMISEAR 215
XX 172 FOYIEGEMRTRIRYNRRS 189
XX 216 FOYIEGEMRTRIRYNRRS 233
XX RESULT 7
XX AAP95639
XX ID AAP95639 standard; protein; 332 AA.
XX AC AAP95639;
XX 25-MAR-2003 (updated)
XX 31-OCT-2002 (updated)

```

DT 13-AUG-1990 (first entry)
 XX Ricin A encoded by insert from plasmid pRA123.
 DE Plasmid pRA123; ricin-A; ricin-B; cytotoxicity.
 XX Ricinus communis.
 OS Synthetic.
 OS
 XX Key Location/Qualifiers
 PH 1...35
 FT /label= leader sequence
 FT Peptide 36..302
 FT /label=A-chain
 FT Peptide 303..314
 FT /label=linker
 FT Peptide 315..332
 FT /label=B-chain
 PN EP335476-A.
 XX
 XX 04-OCT-1989.
 PD
 XX 19-JAN-1989; 89EP-0201162.
 XX
 XX 08-FEB-1984; 84US-0578115.
 PR 08-FEB-1984; 84US-0578121.
 PR 09-FEB-1984; 84US-0578122.
 PR 07-SEP-1984; 84US-0648759.
 PR 20-SEP-1984; 84US-0653515.
 XX
 XX (CETU) CETUS CORPORATION.
 PA
 XX Gelfand D, Lawyer FC, Horn G, Greenfield L, Nitecki D, Kaplan D;
 PI Piatak MJ;
 XX
 DR WPI; 1989-286959/40.
 XX N-PSDB; AAN91281.
 XX
 XX Recombinant vectors expressing ricin chains or diphtheria toxin -used for
 PT prodn. of new immunotoxin conjugates with monoclonal antibodies, having
 PT high cell specificity and good extracellular stability.
 XX
 XX Disclosure; Fig 14; 54pp; English.
 PS
 XX Plasmid RA123 (ATCC No. 39799) carries the entire coding sequence for
 CC ricin A, as well as codons for 12 AAs joining the A to the B chain.
 CC Following modification for ease of manipulation the plasmid was used to
 CC construct expression vectors which express the conjugates in
 CC host cells.
 CC (Updated on 31-OCT-2002 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 XX Sequence 332 AA;
 SQ
 Query Match 97.9%; Score 951; DB 10; Length 332;
 Best Local Similarity 94.9%; Pred. No. 1.7e-93;
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
 QY 2 IFPKQYPIINFITAGATVQSYNTFIRAVRGRLT-----VLPNRYGLPQNFILV 51
 DB 36 IFPKQYPIINFITAGATVQSYNTFIRAVRGRLTGGADVRHEIPVLPNRYGLPQNFILV 95
 QY 52 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEAITHLFTDVQNRYYTFAFG 111
 DB 96 ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEAITHLFTDVQNRYYTFAFG 155
 QY 112 GNYDRLEQAGNLRNIELNGNPLEEALISALYYSTGGTGLPTLARSFFICIQMISEAAR 171
 DB 156 GNYDRLEQAGNLRNIELNGNPLEEALISALYYSTGGTGLPTLARSFFICIQMISEAAR 215

QY 172 FOYIEGEMTRIRYNRRS 189
 DB 216 FOYIEGEMTRIRYNRRS 233
 RESULT 8
 AAR70827
 ID AAR70827 standard; Protein; 554 AA.
 XX
 AC AAR70827;
 XX
 XX 25-MAR-2003 (updated)
 DT 31-AUG-1995 (first entry)
 XX
 XX Anti-cataract immunotoxin.
 XX Immunotoxin; heavy chain; light chain; variable region; antibody;
 KW ricin-A; cytostatic; cataract; lens opacification; epithelial cell;
 KW pHB19; 4197X; monoclonal antibody; Mab.
 XX
 XX Synthetic.
 OS
 XX Location/Qualifiers
 PH Key
 FT Peptide 1..27
 FT /label= Sig_peptide
 FT /note= "phoA signal sequence"
 FT Domain 28..145
 FT /label= HEAVY
 FT /note= "Mab 4197X heavy chain"
 FT Peptide 148..166
 FT /label= LINKER
 FT Domain 169..274
 FT /label= LIGHT
 FT /note= "Mab 419X light chain"
 FT Domain 276..544
 FT /label= RICIN-A
 FT Peptide 549..554
 FT /label= TAG
 FT /note= "hexa-histidine tail"
 XX
 XX WO9503828-A1.
 PN
 XX 09-FEB-1995.
 XX
 XX 15-JUL-1994; 94WO-US07919.
 PF
 XX 02-AUG-1993; 93US-0101329.
 PR
 XX (HOUS-) HOUSTON BIOTECHNOLOGY INC.
 PA
 XX Gould RM, Kelleher PJ, Wallace TL, Wood MS;
 PI WPI; 1995-082036/11.
 XX N-PSDB; AAQ85386.
 DR
 XX New single chain immunotoxin - binds specifically to epithelial
 PT cells, for inhibiting development of sec. cataracts after
 PT extra:capsular cataract extraction.
 PT
 XX Disclosure; Fig.4; 68pp; English.
 PS
 XX The immunotoxin given in AAR70827 comprises the heavy and light chain
 CC variable regions of anti-lens epithelium IgG3 Mab 4197X linked to
 CC ricin-A and a hexa-histidine tag. The DNA construct encoding the
 CC immunotoxin was expressed from pHB19 in E. coli.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 XX Sequence 554 AA;
 SQ
 Query Match 97.9%; Score 951; DB 16; Length 554;
 Best Local Similarity 94.9%; Pred. No. 3.4e-93;
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
 DB 278 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 337
 QY 52 ELSNHAELSVTLALDVTNAYVGVYGRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111
 DB 338 ELSNHAELSVTLALDVTNAYVGVYGRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 397
 QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFFIICQMISEAAR 171
 DB 398 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFFIICQMISEAAR 457
 QY 172 FOYIEGEMTRIRYNRRS 189
 DB 458 FOYIEGEMTRIRYNRRS 475

RESULT 9
 AAP90079 standard; protein; 562 AA.
 XX
 AC AAP90079;
 DT 25-MAR-2003 (updated)
 DT 01-NOV-1989 (first entry)
 DE Ricin D.
 XX
 KW Ricin D; Ricinus communis; castor beans; Zanicbariensis variety;
 KW modified; lectin binding removed; reduced cell binding
 XX
 OS Ricinus communis (castor beans).
 XX
 PN WO8904839-A.
 XX
 PD 01-JUN-1989.
 XX
 PF 23-NOV-1988; 88WO-US04238.
 XX
 PR 24-NOV-1987; 87US-0124735.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Brown EL, Jones S;
 XX
 DR WPI; 1989-178366/24.
 DR N-PSDB; AAN90068.
 XX
 PT Modified ricin molecules and toxin conjugates
 PT - in which the lectin binding function of the B chain
 PT is removed or diminished to reduce cell binding.
 XX
 PS Disclosure; fig 1; 51pp; English.
 XX
 CC Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment
 CC of DNA from Ricinus communis, Zanicbariensis variety. Patent
 CC discloses many modifications of ricin in which the lectin binding
 CC function of the B chain is diminished or removed, and conjugation
 CC to toxins to eliminate cell binding.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 562 AA;
 Query Match 97.9%; Score 951; DB 10; Length 562;
 Best Local Similarity 94.9%; Pred. No. 3.5e-93;
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
 DB 36 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVVRHEIPVLPNRVGLPINQRFILV 95
 QY 52 ELSNHAELSVTLALDVTNAYVGVYGRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111

DB 96 ELSNHAELSVTLALDVTNAYVGVYGRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 155
 QY 112 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFFIICQMISEAAR 171
 DB 156 GNYDRLEQLAGNLRENIELGNGLPLEEASALYYSTGGTQPLTARSFFIICQMISEAAR 215
 QY 172 FOYIEGEMTRIRYNRRS 189
 DB 216 FOYIEGEMTRIRYNRRS 233

RESULT 10
 AAP50166 standard; Protein; 565 AA.
 XX
 AC AAP50166;
 DT 16-OCT-1991 (first entry)
 DE Sequence of preprorin encoded by pRCL617.
 XX
 KW Toxin; anti-tumour therapy.
 XX
 OS Ricinus.
 XX
 PH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= signal
 FT Protein 25..565
 FT Region 292..303
 FT /label= links the C-terminus of the A chain and
 FT the N-terminus of the B chain
 FT Modified-site 34..36
 FT /label= N-linked glycosylation
 FT Modified-site 260..262
 FT /label= N-linked glycosylation
 FT Modified-site 398..400
 FT /label= N-linked glycosylation
 FT Modified-site 438..440
 FT /label= N-linked glycosylation
 XX
 PN EPI45111-A.
 XX
 PD 19-JUN-1985.
 XX
 PF 13-JUL-1984; 84EP-0304801.
 XX
 PR 13-MAR-1984; 84GB-0006569.
 PR 15-JUL-1983; 83GB-0019265.
 PR 15-JUL-1983; 83CH-0019265.
 XX
 PA (UWVA-) UNIV WARWICK.
 XX
 PI Lord JM, Roberts LM, Lamb FI;
 XX
 DR WPI; 1985-148040/25.
 DR N-PSDB; AAN50202.
 XX
 PT New DNA sequences coding for ricin type plant toxin - or its
 PT mutants, and modified vectors and host microorganisms
 XX
 PS Disclosure; Page 30-30c; 40pp; English.
 XX
 CC Preprorin is the whole polypeptide encoded by AAN50202 and the DNA
 CC encoding this is claimed. Preprorin is obtained from preprorin by
 CC removal of the AA leader sequence. The linker AA sequence which is
 CC present in the precursor polypeptide is enzymatically removed in the
 CC cell to separate the A and B chains, which are joined by a
 CC disulphide bridge during the formation of the ricin molecule itself.
 CC This linker region as well as the presumptive amino terminal leader
 CC or signal sequence are not present in the sequences already
 CC published by Funatsu et al.

XX Sequence 565 AA; Query Match 97.9%; Score 951; DB 6; Length 565; Best Local Similarity 94.9%; Pred. No. 3.5e-93; Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IPKQYPIINFITAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILV 51

DB 25 IPKQYPIINFITAGATVQSYNFIKAVRGLTGGADVRHDPVLPNRVGLPINQRFILV 84

QY 52 ELSNHAELSVTLALDVNTAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAG 111

DB 85 ELSNHAELSVTLALDVNTAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAG 144

QY 112 GNYDRLEQLAGNLRENIEIENGPLEEASALYYSTGGTQLPTLARSFFICQMISEAAR 171

DB 145 GNYDRLEQLAGNLRENIEIENGPLEEASALYYSTGGTQLPTLARSFFICQMISEAAR 204

QY 172 FOYIEGEMTRIRYNRRS 189

DB 205 FOYIEGEMTRIRYNRRS 222

RESULT 11

AAG78300

ID AAG78300 standard; Protein; 565 AA.

XX AAG78300;

XX 15-NOV-2001 (first entry)

XX Castor bean preprorin protein (SEQ ID 1).

XX Castor bean plant; preprorin; ricin; A chain; B chain;

KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;

KW retroviral infection; anti-HIV; virucide activity; viral protease.

XX Ricinus communis.

XX Key Location/Qualifiers

FT Peptide 1..24 /label= Signal peptide

FT Protein 25..290 /label= Ricin A chain

FT /note= "N-glycosidase"

FT Peptide 291..302 /label= Linker peptide

FT /note= "Cleaved during activation of ricin"

FT Protein 303..565 /label= Ricin B chain

FT /note= "Galactose/N-acetylgalactosamine-binding lectin"

XX WO200160393-A1.

XX 23-AUG-2001.

XX 15-FEB-2001; 2001WO-US05282.

XX 16-FEB-2000; 2000US-0182759.

XX (BECH-) BECHTEL BWXT IDAHO LLC.

XX Keener WK, Ward TE;

XX WPI; 2001-581908/65.

XX N-PSDB; AAI64137.

XX Novel composition comprising toxin e.g., ricin based antiviral compound

PT useful for treating viral infections such as human immunodeficiency

PT virus infection.

XX Disclosure; Page 47-50; 66pp; English.

XX The sequence relates to preprorin protein encoded by the DNA sequence given in AAI64137. The invention relates to a novel toxin (e.g., ricin) based antiviral agent which is toxic to virus-infected cells, but non-toxic to uninfected cells. The invention has anti-HIV and virucide activities. Its mechanism of action is through inactivation of cellular ribosomes and enhancement of binding of the antiviral agent to galactose residues on cell surfaces, and its cellular internalisation. The invention is useful for treating human immunodeficiency virus infection and other viral infections, especially retroviral infections. The antiviral agent is activated in viral particles or early-stage infected cells, killing the cells upon infection and effectively preventing the integration of the viral genome into the host genome thereby preventing the latency/rebound problem. The agent enters all HIV susceptible cells, and not just cells known to act as host cells for the virus. The antiviral agent remains inert in a cell until degraded in it, unless the cell is infected with the virus, where the viral protease activates it.

XX Sequence 565 AA;

Query Match 97.9%; Score 951; DB 22; Length 565; Best Local Similarity 94.9%; Pred. No. 3.5e-93; Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IPKQYPIINFITAGATVQSYNFIKAVRGLT-----VLPNRVGLPINQRFILV 51

DB 25 IPKQYPIINFITAGATVQSYNFIKAVRGLTGGADVRHDPVLPNRVGLPINQRFILV 84

QY 52 ELSNHAELSVTLALDVNTAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAG 111

DB 85 ELSNHAELSVTLALDVNTAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRVTFAG 144

QY 112 GNYDRLEQLAGNLRENIEIENGPLEEASALYYSTGGTQLPTLARSFFICQMISEAAR 171

DB 145 GNYDRLEQLAGNLRENIEIENGPLEEASALYYSTGGTQLPTLARSFFICQMISEAAR 204

QY 172 FOYIEGEMTRIRYNRRS 189

DB 205 FOYIEGEMTRIRYNRRS 222

RESULT 12

AAG78304

ID AAG78304 standard; Protein; 565 AA.

XX AAG78304;

XX 27-NOV-2001 (first entry)

XX Modified castor bean preprorin (SEQ ID 10).

XX Castor bean plant; preprorin; ricin; A chain; B chain;

KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;

KW retroviral infection; anti-HIV; virucide; viral protease.

XX Chimeric - Ricinus communis

OS Chimeric - Human immunodeficiency virus type 2.

XX Key Location/Qualifiers

FT Peptide 1..24 /label= Signal peptide

FT Protein 25..565 /label= Prorin

FT /note= "Prorin consists of the ricin A chain, a linker peptide, and the ricin B chain. Prorin is proteolytically cleaved between the A chain and the linker to yield mature ricin"

FT Protein 25..291 /label= Ricin A chain

FT /note= "N-glycosidase"

FT Peptide 292..303 /label= Linker peptide

FT Cleavage-site 296..297


```

FT      /label= HIV protease_cleavage_site
FT      304..565
FT      /label= Ricin_B chain
FT      /note= "Galactose/N-acetylgalactosamine-binding lectin"
XX
XX      WO200160393-A1.
XX
XX      23-AUG-2001.
XX
XX      15-FEB-2001; 2001WO-US05282.
XX
XX      16-FEB-2000; 2000US-0182759.
XX      (BECH-) BECHTEL BMYT IDAHO LLC.
XX      Keener WK, Ward TE;
XX
XX      WPI: 2001-581908/65.
XX      N-PSDB; AA164145.
XX
XX      Novel composition comprising toxin e.g., ricin based antiviral compound
XX      useful for treating viral infections such as human immunodeficiency
XX      virus infection.
XX
XX      Example 1; Page 59-63; 66pp; English.
XX
XX      The sequence relates to the amino acid sequence of a modified preproricin
XX      protein encoded by AA164145. The invention relates to a novel toxin
XX      (e.g. ricin) based antiviral agent which is toxic to virus-infected
XX      cells, but non-toxic to uninfected cells. The invention has anti-HIV and
XX      virucide activities. The agent is able to enter all HIV susceptible
XX      cells, and not just cells known to act as host cells for the virus. The
XX      antiviral agent remains inert in a cell unless the cell is infected
XX      with the HIV virus, where the viral protease activates it. Ricin's
XX      mechanism of action is through inactivation of cellular ribosomes and
XX      enhancement of binding of the antiviral agent to galactose residues on
XX      cell surfaces, and its cellular internalization. The invention is useful
XX      for treating human immunodeficiency virus infection and other viral
XX      infections, especially retroviral infections. The antiviral agent is
XX      activated in viral particles or early-stage infected cells, killing the
XX      cells upon infection and effectively preventing the integration of the
XX      viral genome into the host genome thereby preventing the latency/rebound
XX      problem.
XX      Sequence 565 AA;

Query Match      97.9%; Score 951; DB 22; Length 565;
Best Local Similarity 94.9%; Pred. No. 3.5e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY      2  IFPKQYPIINFATTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51
DB      25  IFPKQYPIINFATTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 84
QY      52  ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 111
DB      85  ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 144
QY      112 GNYDRLEQLAGNLRENIELGNGFLEAISAALYYSTGGTQLPTLARSFFICIMISEAR 171
DB      145 GNYDRLEQLAGNLRENIELGNGFLEAISAALYYSTGGTQLPTLARSFFICIMISEAR 204
QY      172 FOVIEGEMTRIRNRRS 189
DB      205 FOVIEGEMTRIRNRRS 222

RESULT 13
AAP70326
ID      AAP70326 standard; Protein; 576 AA.
XX
XX      AAP70326;
XX

```

```

DT      25-MAR-2003 (updated)
DT      21-MAY-1991 (first entry)
XX
XX      Sequence of Ricinus communis (castor bean) Ricin toxin
XX      (RT or ricin) E precursor encoded by prt38.
XX
XX      Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
XX      plant toxin.
XX
XX      Ricinus communis.
XX
XX      Key      Location/Qualifiers
XX      Region 1..35
XX      /note= "leader"
XX      Region 36..302
XX      /note= "A-chain"
XX      Region 315..576
XX      /note= "B-chain"
XX
XX      EP237676-A.
XX
XX      23-SEP-1987.
XX
XX      13-NOV-1986; 86EP-0308877.
XX
XX      07-MAR-1986; 86US-0837583.
XX      (CETU ) CETUS CORP.
XX      (CHIR ) CHIRON CORP.
XX
XX      Piatak M;
XX
XX      WPI: 1987-265177/38.
XX      N-PSDB; AA70526.
XX
XX      New non-glycosylated ricin precursor and toxin etc. - are prepd.
XX      by recombinant DNA procedures with specific isolation steps for
XX      purer and soluble prods.
XX
XX      Disclosure; Fig 14(1-2); 112pp; English.
XX
XX      The full length sequences encoding ricin A (AA70520), ricin D
XX      (AA70525) putative ricin E (AA70526) and RCA (AA70524) in precursor
XX      form were obtained, using the messenger RNA to obtain a cDNA library, and
XX      then probing the library to retrieve the desired cDNA inserts. The
XX      library was probed using the 35-mer given in AA70514. Figure 4 (see
XX      AA70520, AA70521, AA70522), shows the nucleotide sequences of three
XX      plasmids containing cDNA inserts obtained by probing a cDNA library
XX      for sequences encoding ricin B using the probe in AA70517. The cDNA
XX      inserts can be placed into expression vectors. Site-directed
XX      mutagenesis may be used to place an ATG start codon and a HindIII
XX      site at the beginning of the mature protein. (see AA70518). The
XX      coding sequences of the inserts can be ligated into expression
XX      vectors containing the Ph0A promoter-operator and leader sequence
XX      (AA70523) and suitable retroregulators.
XX      (Updated on 25-MAR-2003 to correct PA field.)
XX      Sequence 576 AA;

Query Match      97.9%; Score 951; DB 8; Length 576;
Best Local Similarity 94.9%; Pred. No. 3.6e-93;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY      2  IFPKQYPIINFATTAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINQRFILV 51
DB      36  IFPKQYPIINFATTAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINQRFILV 95
QY      52  ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 111
DB      96  ELSNHAELSVTLLDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRYYTFAFG 155
QY      112 GNYDRLEQLAGNLRENIELGNGFLEAISAALYYSTGGTQLPTLARSFFICIMISEAR 171

```

DB 156 GNYDRLEQLAGNLELNGPLEEAISALYYSTGGTQTLPTLARSFIICQMISEAAR 215

QY 172 FQYIEGEMRTIRYNRRS 189
 DB 216 FQYIEGEMRTIRYNRRS 233

RESULT 14
 AAW25787
 ID AAW25787 standard; Protein; 576 AA.

XX AC AAW25787;
 XX XX
 DT 25-MAR-2003 (updated)
 DT 27-MAR-1998 (first entry)
 XX XX
 DE Castor bean ricin.
 XX XX
 KW Ricin; cytotoxin; hybrid protein; cell delivery;
 KW cell binding ligand; translocation domain; diphtheria toxin B';
 KW interleukin-2; T-cell lymphoma; organ rejection; therapy.
 XX OS
 XX Ricinus communis.

XX Key Location/Qualifiers
 FT Peptide 1..35
 FT /label= Sig_peptide
 FT Protein 36..302
 FT /label= A-domain
 FT Peptide 303..314
 FT /label= Linker
 FT Domain 315..576
 FT /label= B-domain

XX US5668255-A.
 XX 16-SEP-1997.
 XX 04-AUG-1993; 93US-0102387.
 XX 27-JUN-1991; 91US-0722484.
 PR 07-JUN-1984; 84US-0618199.
 PR 25-APR-1985; 85US-0726808.
 PR 07-JUN-1985; 85US-0742554.
 PR 22-DEC-1989; 89US-0456095.
 PR 14-JUN-1990; 90US-0538276.
 PR 04-AUG-1993; 93US-0102387.
 XX (SERA-) SERAGEN INC.
 PA
 XX Murphy JR;
 XX WPI; 1997-470103/43.
 DR N-PSDB; AAT91638.
 XX
 PT New hybrid molecules for delivery of agents to cells - comprise a
 PT binding domain of a cell binding ligand and a portion of a
 PT translocation domain of a protein
 XX
 PS Example 4; Fig 11A-B; 30pp; English.
 XX
 CC This polypeptide comprises the castorbean cytotoxin, ricin.
 CC DNA (see AAT91638) encoding the enzymatic A domain and a portion
 CC of the A-to-B linker peptide of ricin was used to construct a
 CC ricin-diphtheria toxin B'-interleukin-2 gene that was expressed in
 CC E. coli. The hybrid protein can be isolated and used to treat
 CC conditions involving over-production of cells bearing IL2 receptors,
 CC such as certain T-cell lymphomas and organ transplant rejection
 CC crises. The hybrid inactivates ribosomes in cells bearing IL2
 CC receptors, resulting in cessation of protein synthesis and death of
 CC target cells. Claimed hybrid proteins comprise a translocation
 CC domain and a cell binding domain from e.g. a hormone, growth factor
 CC or polypeptide toxin. The hybrid molecules can be used for the

CC delivery of agents (e.g. therapeutic genes, toxins, detectable
 CC labels) into cells. The use of a translocation mechanism ensures
 CC that the hybrid will be effective in relatively low doses, since a
 CC high proportion of the substance of interest will be taken into the
 CC targeted cells. The hybrid molecules can be manufactured as a
 CC single hybrid recombinant protein, permitting reproducibility,
 CC consistency, and the precise control of composition.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX XX

SQ Sequence 576 AA;
 Query Match 97.9%; Score 951; DB 18; Length 576;
 Best Local Similarity 94.9%; Pred No. 3.6e-93;
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFETTAGTAGVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
 DB 36 IFPKQYPIINFETTAGTAGVQSYTNFIRAVRGRLTGDVDRHEIPVLPNRVGLPINQRFILV 95

QY 52 ELSNHAELSVTALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHLFTDVQNYTFAFG 111
 DB 96 ELSNHAELSVTALDVTNAYVGYRAGNSAYFFHPDNOEDAEATHLFTDVQNYTFAFG 155

QY 112 GNYDRLEQLAGNLELNGPLEEAISALYYSTGGTQTLPTLARSFIICQMISEAAR 171
 DB 156 GNYDRLEQLAGNLELNGPLEEAISALYYSTGGTQTLPTLARSFIICQMISEAAR 215

QY 172 FQYIEGEMRTIRYNRRS 189
 DB 216 FQYIEGEMRTIRYNRRS 233

RESULT 15
 AAY55892
 ID AAY55892 standard; Protein; 576 AA.
 XX AC AAY55892;
 XX XX
 DT 15-FEB-2000 (first entry)
 XX XX
 DE Castor bean ricin toxin.
 XX XX
 KW Recombinant; hybrid; binding domain; ligand; animal cell; diphtheria;
 KW translocation domain; botulinum; neurotoxin; ricin; cholera; tetanus;
 KW shiga-like toxin; pertussis; translocation; cytoplasmic membrane; HIV;
 KW cytosol; therapy; genetic deficiency disease; enzyme; co-factor; poison;
 KW adipocyte; cancer; virus; infection; antibody.
 XX OS
 XX Ricinus communis.
 XX US5965406-A.
 XX 12-OCT-1999.
 XX 07-JUN-1995; 95US-0488246.
 XX 04-AUG-1993; 93US-0102387.
 PR 07-JUN-1984; 84US-0618199.
 PR 27-JUN-1991; 91US-0722484.
 PR 25-APR-1985; 85US-0726808.
 PR 07-JUN-1985; 85US-0742554.
 PR 22-DEC-1989; 89US-0456095.
 PR 14-JUN-1990; 90US-0538276.
 XX (SERA-) SERAGEN INC.
 PA
 XX Murphy JR;
 XX WPI; 1999-632431/54.
 DR N-PSDB; AAZ30663.
 XX Recombinant DNA molecule encoding a three part hybrid protein used in
 PT the treatment of Aids and genetic deficiency diseases -

XX

PS Example 4; Fig 11; 31pp; English.

XX

CC The invention relates to a recombinant DNA molecule encoding a hybrid
CC protein comprising three parts: (a) the first part comprises a portion
CC of the binding domain of a cell-binding polypeptide ligand allowing the
CC hybrid protein to bind to an animal cell; (b) the second part comprises
CC a portion of a translocation domain of a naturally occurring protein
CC selected from diphtheria toxin, botulinum neurotoxin, ricin, cholera
CC toxin, LT toxin, C3 toxin, shiga-like toxin, pertussis toxin and tetanus
CC toxin, which translocate the third part of the across the cytoplasmic
CC membrane into the cytosol of the cell; and (c) the third part comprises
CC a polypeptide entity to be introduced into the cell, which is non-native
CC to the naturally occurring protein of (b). This sequence represents the
CC Castor bean ricin toxin sequence for use in generating the hybrid of the
CC invention. The hybrid molecule enables the direction of appropriate
CC therapy to affected cells, allowing them to function properly and
CC alleviate or cure the disease. The hybrid is especially used in treating
CC genetic deficiency diseases, by delivering to affected cells an enzyme
CC supplying the missing function, to supplementing cellular levels of a
CC particular enzyme or a scarce precursor or cofactor, to directing toxins
CC or other poisons to destroy particular cells (such as adipocytes, cancer
CC cell, or virus infected-cells), to counteracting viral infections such as
CC HIV, by introducing appropriate antibodies to viral proteins. It is also
CC involved in the process of getting non-therapeutic substances such as
CC detectable labels into cells.

XX

SQ Sequence 576 AA;

Query Match	97.9%;	Score 951;	DB 20;	Length 576;
Best Local Similarity	94.9%;	Pred. No. 3.6e-93;		
Matches 188;	Conservative 0;	Mismatches 0;	Indels 10;	Gaps 1;
QY	2	IFPKQYPIINFITAGATVQSYTNFIRAVRGLT-----VLPNRVGLPINOQRFILV	51	
Db	36	IFPKQYPIINFITAGATVQSYTNFIRAVRGLTGGADVRHEIPVLPNRVGLPINOQRFILV	95	
QY	52	ELSNHAELSVTLALDVNTNAYVVGVRAGNSGAYFFHPDNOQEDAEAIHTLFTDVQNRYYTFAFG	111	
Db	96	ELSNHAELSVTLALDVNTNAYVVGVRAGNSGAYFFHPDNOQEDAEAIHTLFTDVQNRYYTFAFG	155	
QY	112	GNVDRLEQLAGNLRNIELGNGLPEEAISALYYISTGGTOLPTLARSFIIQIMISEAAR	171	
Db	156	GNVDRLEQLAGNLRNIELGNGLPEEAISALYYISTGGTOLPTLARSFIIQIMISEAAR	215	
QY	172	FOYIEGEMRTRIRYNRRS	189	
Db	216	FOYIEGEMRTRIRYNRRS	233	

Search completed: February 10, 2004, 16:22:27

Job time : 33.2239 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 10.9281 Seconds
(without alignments)
731.761 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971

Sequence: 1 MIFPKQPIINFTAGATVQ.....ARFQVIEGEMTRIRYNRRS 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	956	98.5	268	2	US-08-356-786-8
2	956	98.5	534	2	US-08-356-786-10
3	951	97.9	267	1	US-07-901-707-1
4	951	97.9	267	1	US-07-988-430-1
5	951	97.9	267	1	US-08-425-336-1
6	951	97.9	267	1	US-08-488-1138-1
7	951	97.9	267	1	US-08-477-484B-1
8	951	97.9	267	2	US-08-646-360-1
9	951	97.9	267	3	US-08-839-765-1
10	951	97.9	267	3	US-09-136-389-1
11	951	97.9	267	4	US-09-610-838-1
12	951	97.9	267	5	PCT-US92-09487-1
13	951	97.9	290	1	US-08-378-761A-27
14	951	97.9	290	1	US-08-485-286-27
15	951	97.9	290	6	5248606-4
16	941	96.9	267	1	US-08-218-303-16
17	941	96.9	267	2	US-08-338-793D-61
18	941	96.9	267	4	US-09-538-873-1
19	861.5	88.7	540	1	US-08-378-761A-77
20	861.5	88.7	540	1	US-08-485-286-77
21	336	34.6	247	1	US-08-488-113B-6
22	336	34.6	247	1	US-08-477-484B-6
23	336	34.6	247	2	US-08-646-360-6
24	336	34.6	247	3	US-08-839-765-6
25	336	34.6	247	3	US-09-136-389-6
26	336	34.6	247	4	US-09-610-838-6
27	336	34.6	267	1	US-08-378-761A-74

28 336 34.6 267 1 US-08-485-286-74 Sequence 74, Appl
29 336 34.6 289 1 US-07-923-692C-4 Sequence 4, Appl
30 336 34.6 289 1 US-08-184-237-4 Sequence 4, Appl
31 336 34.6 289 2 US-08-482-920-4 Sequence 4, Appl
32 336 34.6 289 3 US-08-484-341-4 Sequence 4, Appl
33 336 34.6 289 3 US-08-483-502-4 Sequence 4, Appl
34 336 34.6 289 4 US-09-726-651A-4 Sequence 4, Appl
35 320.5 33.0 282 1 US-08-324-301-15 Sequence 15, Appl
36 310.5 32.0 255 1 US-07-901-707-6 Sequence 6, Appl
37 310.5 32.0 255 1 US-07-988-430-6 Sequence 6, Appl
38 310.5 32.0 255 1 US-08-425-336-6 Sequence 6, Appl
39 310.5 32.0 255 5 PCT-US92-09487-6 Sequence 6, Appl
40 305 31.4 248 3 US-08-902-486-7 Sequence 7, Appl
41 305 31.4 290 1 US-08-245-754A-2 Sequence 2, Appl
42 305 31.4 290 2 US-08-597-731-2 Sequence 2, Appl
43 305 31.4 496 3 US-08-902-486-15 Sequence 15, Appl
44 303.5 31.3 250 1 US-08-378-761A-71 Sequence 71, Appl
45 303.5 31.3 250 1 US-08-485-286-71 Sequence 71, Appl

ALIGNMENTS

RESULT 1
US-08-356-786-8
; Sequence 8, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; TITLE OF INVENTION: Marker
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/356,786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-786-8

Query Match 98.5%; Score 956; DB 2; Length 268;
Best Local Similarity 95.0%; Pred. NO. 7.1e-103;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50
Db 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 60
QY 51 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTF 110
Db 61 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTF 120
QY 111 GGNDRLEQLAGNLRNENIELGNGLPLEEALISALYYSTGGTQPLTLARSFIICIMISEAA 170
Db 121 GGNDRLEQLAGNLRNENIELGNGLPLEEALISALYYSTGGTQPLTLARSFIICIMISEAA 180
QY 171 RFQYIEGEMRTRIRYNRS 189
Db 181 RFQYIEGEMRTRIRYNRS 199

RESULT 2

US-08-356-786-10
; Sequence 10, Application US/08356786
; Patent No. 5877305
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer
; MARKER
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08356786
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/831,967
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitcher, Edmund R.
; REGISTRATION NUMBER: 27,829
; REFERENCE/DOCKET NUMBER: CRP-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 248-7000
; TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 534 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-356-786-10

Query Match 98.5%; Score 956; DB 2; Length 534;
Best Local Similarity 95.0%; Pred. No. 2e-102;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50
Db 3 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 62
QY 51 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTF 110

Db 63 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTF 122
QY 111 GGNDRLEQLAGNLRNENIELGNGLPLEEALISALYYSTGGTQPLTLARSFIICIMISEAA 170
Db 123 GGNDRLEQLAGNLRNENIELGNGLPLEEALISALYYSTGGTQPLTLARSFIICIMISEAA 182
QY 171 RFQYIEGEMRTRIRYNRS 189
Db 183 RFQYIEGEMRTRIRYNRS 201

RESULT 3

US-07-901-707-1
; Sequence 1, Application US/07901707
; Patent No. 5376546
; GENERAL INFORMATION:
; APPLICANT: Bernhardt, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Steve F.
; APPLICANT: Lane, Julie A.
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell,
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07901,707
; FILING DATE: 19920619
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5376546and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27129/30910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-5750
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-901-707-1

Query Match 97.9%; Score 951; DB 1; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 51
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 60
QY 52 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTF 111
Db 61 VELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNYTF 120
QY 112 GGNDRLEQLAGNLRNENIELGNGLPLEEALISALYYSTGGTQPLTLARSFIICIMISEAA 171

Db 121 GNYDRLEQAGNLRNIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAAR 180
QY 172 FOYIEGEMTRIRYNRS 189
Db 181 FOYIEGEMTRIRYNRS 198

RESULT 4
US-07-988-430-1
; Sequence 1, Application US/07988430
; Patent No. 5416202
; GENERAL INFORMATION:
; APPLICANT: Bernhard, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Lane, Julie A.
; APPLICANT: Lei, Shau-Ping
; TITLE OF INVENTION: Materials Comprising and Methods of
; TITLE OF INVENTION: Preparation and Use for Ribosome-Inactivating Proteins
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/988,430
; FILING DATE: 19921209
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5416202and, Greta E.
; REGISTRATION NUMBER: 35302
; REFERENCE/DOCKET NUMBER: 31133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-988-430-1

Query Match 97.9%; Score 951; DB 1; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFLV 51
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINORFLV 60
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 111
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 120

QY 112 GNYDRLEQAGNLRNIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAAR 171
Db 121 GNYDRLEQAGNLRNIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAAR 180
QY 172 FOYIEGEMTRIRYNRS 189
Db 181 FOYIEGEMTRIRYNRS 198

RESULT 5
US-08-425-336-1
; Sequence 1, Application US/08425336
; Patent No. 5621083
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; ADDRESSEE: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,336
; FILING DATE: 18-APR-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Thomas C.
; REGISTRATION NUMBER: P-36,989
; REFERENCE/DOCKET NUMBER: 31394
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-425-336-1

Query Match 97.9%; Score 951; DB 1; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 186; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFLV 51
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRRHEIPVLPNRVGLPINORFLV 60
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 111
Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 120
QY 112 GNYDRLEQAGNLRNIELNGNPLEEASALYYSTGGTQPLTARSFIICQMISEAAR 171

Db 121 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGGTQPLTLARSLFIQIMISEAAR 180

QY 172 FOYIEGEMTRIRYNRRS 189

Db 181 FOYIEGEMTRIRYNRRS 198

RESULT 6
US-08-488-113B-1
; Sequence 1, Application US/08488113B
; Patent No. 5744580
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,113B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-113B-1

Query Match 97.9%; Score 951; DB 1; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;
QY 2 IFPKQYPIINFETAGATQSYTNFIRAVRGRIT-----VLPNRVGLPINQRFILV 51
Db 1 IFPKQYPIINFETAGATQSYTNFIRAVRGRIT-----VLPNRVGLPINQRFILV 60

QY 52 ELSNHAELSVTLALDVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDYQNRVTFAPG 111

Db 61 ELSNHAELSVTLALDVTNAYVYVGRAGNSAYFFHPDNOEDAEATHLFTDYQNRVTFAPG 120

QY 112 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGGTQPLTLARSLFIQIMISEAAR 171

Db 121 GNYDRLEQLAGNLRNIELGNGLPLEEALISALYYSTGGTQPLTLARSLFIQIMISEAAR 180

QY 172 FOYIEGEMTRIRYNRRS 189

Db 181 FOYIEGEMTRIRYNRRS 198

RESULT 7
US-08-477-484B-1
; Sequence 1, Application US/08477484B
; Patent No. 5756699
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,484B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-477-484B-1

Query Match 97.9%; Score 951; DB 1; Length 267;

Best Local Similarity 94.9%; Pred. No. 2.7e-102; Indels 10; Gaps 1;
Matches 188; Conservative 0; Mismatches 0;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVDRHEIPVLPNRVGLPINQRFILV 60
QY 52 ELSNHAELSVTLLDVTNAYVGVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111
Db 61 ELSNHAELSVTLLDVTNAYVGVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 120
QY 112 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYVSTGTQTLPTLARSFIIICIMISEAAR 171
Db 121 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYVSTGTQTLPTLARSFIIICIMISEAAR 180
QY 172 FOYIEGEMRTRIRYNNRS 189
Db 181 FOYIEGEMRTRIRYNNRS 198

RESULT 8
US-08-646-360-1
; Sequence 1, Application US/08646360
; Patent No. 5837491
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-646-360-1

Query Match 97.9%; Score 951; DB 2; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102; Indels 10; Gaps 1;
Matches 188; Conservative 0; Mismatches 0;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGDVDRHEIPVLPNRVGLPINQRFILV 60
QY 52 ELSNHAELSVTLLDVTNAYVGVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 111
Db 61 ELSNHAELSVTLLDVTNAYVGVGRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFG 120
QY 112 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYVSTGTQTLPTLARSFIIICIMISEAAR 171
Db 121 GNYDRLEQLAGNLRENIELGNGPLEEALISALYYVSTGTQTLPTLARSFIIICIMISEAAR 180
QY 172 FOYIEGEMRTRIRYNNRS 189
Db 181 FOYIEGEMRTRIRYNNRS 198

RESULT 9
US-08-839-765-1
; Sequence 1, Application US/08839765
; Patent No. 6146631
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/839,765
; FILING DATE: 15-APR-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/425,336
; FILING DATE: 18-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3
; TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-839-765-1

Query Match          97.9%; Score 951; DB 3; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY      2  IFPKQYPIINFITAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db      1  IFPKQYPIINFITAGATVQSYTNFIRAVRGRLTGADVREHVPVLPNRVGLPINQRFILV 60

QY      52  ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111
Db      61  ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 120

QY      112  GNYDRLEQLAGNLRNLELGNGLPEEAISALYYSTGCTQLPTLARSFIICIMISEAAR 171
Db      121  GNYDRLEQLAGNLRNLELGNGLPEEAISALYYSTGCTQLPTLARSFIICIMISEAAR 180

QY      172  FOYIEGEMRTRIRYNRS 189
Db      181  FOYIEGEMRTRIRYNRS 198

RESULT 10
US-09-136-389-1
; Sequence 1, Application US/09136389
; Patent No. 6146850
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/136,389
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-136-389-1

Query Match          97.9%; Score 951; DB 3; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY      2  IFPKQYPIINFITAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db      1  IFPKQYPIINFITAGATVQSYTNFIRAVRGRLTGADVREHVPVLPNRVGLPINQRFILV 60

QY      52  ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 111
Db      61  ELSNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFG 120

QY      112  GNYDRLEQLAGNLRNLELGNGLPEEAISALYYSTGCTQLPTLARSFIICIMISEAAR 171
Db      121  GNYDRLEQLAGNLRNLELGNGLPEEAISALYYSTGCTQLPTLARSFIICIMISEAAR 180

QY      172  FOYIEGEMRTRIRYNRS 189
Db      181  FOYIEGEMRTRIRYNRS 198

RESULT 11
US-09-610-838-1
; Sequence 1, Application US/09610838
; Patent No. 6376217
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; TITLE OF INVENTION: Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,838
; FILING DATE: 06-JUL-2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/136,389
; FILING DATE: 18-AUG-1998
; APPLICATION NUMBER: 08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
```

```

; FILING DATE: 12-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/707-8889
; TELEFAX: 312/707-9155
; TELEX: 650 388-1248
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-610-838-1

```

```

Query Match      97.9%; Score 951; DB 4; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

Qy  2  IFKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db  1  IFKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVHREIPVLPNRVGLPINQRFILV 60

Qy  52  ELSNHAELSVTLADVNAVVGVRAGNSAYFFHPDNOEDAEALTHLFTDVQNYTFAFG 111
Db  61  ELSNHAELSVTLADVNAVVGVRAGNSAYFFHPDNOEDAEALTHLFTDVQNYTFAFG 120

Qy  112  GNYDRLEQLAGNLENIELGNGLPLEEAISALYYSTGTQTLPTLARSFIIQIMISEAAR 171
Db  121  GNYDRLEQLAGNLENIELGNGLPLEEAISALYYSTGTQTLPTLARSFIIQIMISEAAR 180

Qy  172  FQYIEGEMRTRIRYNRRS 189
Db  181  FQYIEGEMRTRIRYNRRS 198

```

```

RESULT 12
PCT-US92-09487-1
; Sequence 1, Application PC/TUS9209487
; GENERAL INFORMATION:
; APPLICANT: Bernhardt, Susan L.
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Lane, Julie A.
; APPLICANT: Lei, Shau-Ping
; TITLE OF INVENTION: Materials Comprising and Methods of
; PREPARATION AND USE FOR RIBOSOME-INACTIVATING PROTEINS
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Hicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/09487
; FILING DATE: 19921104
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Noland, Greta E.
; REGISTRATION NUMBER: 35302
; REFERENCE/DOCKET NUMBER: 31133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-09487-1

```

```

Query Match      97.9%; Score 951; DB 5; Length 267;
Best Local Similarity 94.9%; Pred. No. 2.7e-102;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

Qy  2  IFKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db  1  IFKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVHREIPVLPNRVGLPINQRFILV 60

Qy  52  ELSNHAELSVTLADVNAVVGVRAGNSAYFFHPDNOEDAEALTHLFTDVQNYTFAFG 111
Db  61  ELSNHAELSVTLADVNAVVGVRAGNSAYFFHPDNOEDAEALTHLFTDVQNYTFAFG 120

Qy  112  GNYDRLEQLAGNLENIELGNGLPLEEAISALYYSTGTQTLPTLARSFIIQIMISEAAR 171
Db  121  GNYDRLEQLAGNLENIELGNGLPLEEAISALYYSTGTQTLPTLARSFIIQIMISEAAR 180

Qy  172  FQYIEGEMRTRIRYNRRS 189
Db  181  FQYIEGEMRTRIRYNRRS 198

```

```

RESULT 13
US-08-378-761A-27
; Sequence 27, Application US/08378761A
; Patent No. 5635384
; GENERAL INFORMATION:
; APPLICANT: WALSH, TERENCE A
; APPLICANT: HEY, TIMOTHY D
; APPLICANT: MORGAN, ALICE ER
; TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE
; PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF
; USING
; NUMBER OF SEQUENCES: 81
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ANDREA T. BORUCKI
; STREET: 9330 ZIONSVILLE ROAD
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: US
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/378,761A

FILING DATE: 26-JAN-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: BORUCKI, ANDREA T

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 38272B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 337-4846

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 290 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-378-761A-27

Query Match 97.9%; Score 951; DB 1; Length 290;

Best Local Similarity 94.9%; Pred. No. 3e-102;

Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```
QY      2  IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db      25  IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFILV 84

QY      52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 111
Db      85  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 144

QY     112  GNYDRLEQLAGNLRNIELNGPLERAI SALLYSTGGTQPLTLARSFIICQMISEAAR 171
Db     145  GNYDRLEQLAGNLRNIELNGPLERAI SALLYSTGGTQPLTLARSFIICQMISEAAR 204

QY     172  FOYIEGEMRTRIRYNRRS 189
Db     205  FOYIEGEMRTRIRYNRRS 222
```

RESULT 14

US-08-485-286-27

Sequence 27, Application US/08485286

Patent No. 5646026

Patent No. 5646026 5646119

GENERAL INFORMATION:

APPLICANT: WALSH, TERENCE A

APPLICANT: HEY, TIMOTHY D

APPLICANT: MORGAN, ALICE ER

TITLE OF INVENTION: RIBOSOME-INACTIVATING PROTEINS, INACTIVE

TITLE OF INVENTION: PRECURSOR FORMS THEREOF, A PROCESS FOR MAKING A METHOD OF

NUMBER OF SEQUENCES: 81

CORRESPONDENCE ADDRESS:

ADDRESSEE: ANDREA T. BORUCKI

STREET: 9330 ZIONSVILLE ROAD

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: US

ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,286

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/378761

FILING DATE: 26-JAN-1995

ATTORNEY/AGENT INFORMATION:

NAME: BORUCKI, ANDREA T

REGISTRATION NUMBER: 33651

REFERENCE/DOCKET NUMBER: 38272B

TELECOMMUNICATION INFORMATION:

TELEPHONE: (317) 337-4846

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 290 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-485-286-27

Query Match 97.9%; Score 951; DB 1; Length 290;

Best Local Similarity 94.9%; Pred. No. 3e-102;

Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```
QY      2  IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db      25  IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFILV 84

QY      52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 111
Db      85  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 144

QY     112  GNYDRLEQLAGNLRNIELNGPLERAI SALLYSTGGTQPLTLARSFIICQMISEAAR 171
Db     145  GNYDRLEQLAGNLRNIELNGPLERAI SALLYSTGGTQPLTLARSFIICQMISEAAR 204

QY     172  FOYIEGEMRTRIRYNRRS 189
Db     205  FOYIEGEMRTRIRYNRRS 222
```

RESULT 15

5248606-4

Patent No. 5248606

APPLICANT: WALSH, TERENCE A; HEY, TIMOTHY D; MORGAN,

ALICE E.R.

TITLE OF INVENTION: DNA ENCODING INACTIVE PRECURSOR AND

ACTIVE FORMS OF MAIZE RIBOSOME INACTIVATION

NUMBER OF SEQUENCES: 49

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/535,636

FILING DATE: 11-JUN-1990

SEQ ID NO:4

LENGTH: 290

5248606-4

Query Match 97.9%; Score 951; DB 6; Length 290;

Best Local Similarity 94.9%; Pred. No. 3e-102;

Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

```
QY      2  IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
Db      25  IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFILV 84

QY      52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 111
Db      85  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRYYTFAFG 144

QY     112  GNYDRLEQLAGNLRNIELNGPLERAI SALLYSTGGTQPLTLARSFIICQMISEAAR 171
Db     145  GNYDRLEQLAGNLRNIELNGPLERAI SALLYSTGGTQPLTLARSFIICQMISEAAR 204

QY     172  FOYIEGEMRTRIRYNRRS 189
Db     205  FOYIEGEMRTRIRYNRRS 222
```

Search completed: February 10, 2004, 16:29:32

Sun Feb 15 07:30:00 2004

us-10-083-336a-6.rai

Page 9

Job time : 10.9281 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:26:46 ; Search time 24.3781 Seconds
(without alignment)
1623.314 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971

Sequence: 1 MIFPKQYPIINFTTAGATVQ.....ARFQYIEGEMTRIVNRS 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	971	100.0	189	12	US-10-083-336A-6
2	971	100.0	190	12	US-10-083-336A-11
3	966	99.5	188	12	US-10-083-336A-4
4	956	98.5	188	12	US-10-083-336A-8
5	956	98.5	199	12	US-10-083-336A-5
6	956	98.5	200	12	US-10-083-336A-10
7	951	97.9	198	12	US-10-083-336A-3
8	951	97.9	267	12	US-10-127-890-1
9	951	97.9	576	12	US-10-083-336A-1
10	941	96.9	198	12	US-10-083-336A-7
11	941	96.9	267	12	US-10-282-935-1
12	941	96.9	267	12	US-10-440-796-1
13	939.5	96.8	185	12	US-10-083-336A-9
14	679	69.9	179	12	US-10-083-336A-2
15	336	34.6	247	10	US-09-792-793A-39

16	336	34.6	247	12	US-10-127-890-6
17	336	34.6	247	12	US-10-375-209A-39
18	336	34.6	289	12	US-10-280-679B-4
19	305	31.4	247	10	US-09-792-793A-34
20	305	31.4	247	12	US-10-375-209A-34
21	297.5	30.6	251	12	US-10-282-935-3
22	297.5	30.6	251	12	US-10-440-796-3
23	273	28.1	263	12	US-10-127-890-7
24	272	28.0	263	12	US-10-127-890-4
25	257.5	26.2	248	12	US-10-127-890-5
26	254	26.2	252	9	US-09-347-064-2
27	254	26.2	252	9	US-09-347-064-8
28	245.5	25.3	251	12	US-10-127-890-107
29	244.5	25.2	251	12	US-10-127-890-106
30	244.5	25.2	251	12	US-10-127-890-110
31	244.5	25.2	251	12	US-10-127-890-111
32	243.5	25.1	251	9	US-09-765-527-247
33	243.5	25.1	251	12	US-10-127-890-2
34	243.5	25.1	251	12	US-10-127-890-99
35	243.5	25.1	251	12	US-10-127-890-100
36	243.5	25.1	251	12	US-10-127-890-101
37	243.5	25.1	251	12	US-10-127-890-102
38	243.5	25.1	251	12	US-10-127-890-103
39	243.5	25.1	251	12	US-10-127-890-104
40	243.5	25.1	251	12	US-10-127-890-105
41	243.5	25.1	316	12	US-10-074-596-1
42	243.5	25.1	507	12	US-10-074-596-11
43	242.5	25.0	251	12	US-10-127-890-109
44	242.5	25.0	293	9	US-09-765-527-259
45	242.5	25.0	309	9	US-09-765-527-253

ALIGNMENTS

RESULT 1

US-10-083-336A-6

; Sequence 6, Application US/10083336A

; Publication No. US20030181665A1

; GENERAL INFORMATION:

; APPLICANT: Olson, Mark A

; APPLICANT: Millard, Charles B

; APPLICANT: Byrne, Michael P

; APPLICANT: Wamemacher, Robert W

; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof

; FILE REFERENCE: P6745200 (RIID 01-58)

; CURRENT APPLICATION NUMBER: US/10/083,336A

; CURRENT FILING DATE: 2002-05-21

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 189

; TYPE: PRT

; ORGANISM: Ricinus communis

; US-10-083-336A-6

Query Match	100.0%	Score 971;	DB 12;	Length 189;
Best Local Similarity	100.0%	Pred. No. 2.7e-103;		
Matches 189;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTVLPNRVGLPINQRFILVELSNHAELS	60	
Db	1	MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTVLPNRVGLPINQRFILVELSNHAELS	60	
QY	61	VTALDVTNAYVVGVRAGNSAYFFPHDQDEAETHLFTDVQNYTFAFGNYDRLEQL	120	
Db	61	VTALDVTNAYVVGVRAGNSAYFFPHDQDEAETHLFTDVQNYTFAFGNYDRLEQL	120	
QY	121	AGNLRENIELGNPLEEASALYYSTGTQTLPTLARSFIIICIMTSEARFOYIEGEMR	180	
Db	121	AGNLRENIELGNPLEEASALYYSTGTQTLPTLARSFIIICIMTSEARFOYIEGEMR	180	
QY	181	TRIVNRS	189	

```
Db 181 TRIRYNRS 189
|||||
Db 181 TRIRYNRS 189

RESULT 2
US-10-083-336A-11
; Sequence 11, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnie, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 11
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-11

Query Match 100.0%; Score 971; DB 12; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.7e-103;
Matches 189; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MIPKQYPIINFTTAGATVQSYTNFIRAVRGRLTLPNVRVGLPINQRFILVELSNHAELS 60
Db 1 MIPKQYPIINFTTAGATVQSYTNFIRAVRGRLTLPNVRVGLPINQRFILVELSNHAELS 60
Qy 61 VTALDVTNAYVVGVRAGNSAYFFHFDNQEDAEAIHTLFTDVQNRVYTFAGGNYDRLEQL 120
Db 61 VTALDVTNAYVVGVRAGNSAYFFHFDNQEDAEAIHTLFTDVQNRVYTFAGGNYDRLEQL 120
Qy 121 AGNREINELGNGLPBEAISALYYSTGGTQPLTLARSFFIICIMISEAARFOYIEGEMR 180
Db 121 AGNREINELGNGLPBEAISALYYSTGGTQPLTLARSFFIICIMISEAARFOYIEGEMR 180
Qy 181 TRIRYNRS 189
Db 181 TRIRYNRS 189

RESULT 3
US-10-083-336A-4
; Sequence 4, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnie, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-4

Query Match 99.5%; Score 966; DB 12; Length 188;
Best Local Similarity 100.0%; Pred. No. 9.9e-103;
Matches 188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 IPPKQYPIINFTTAGATVQSYTNFIRAVRGRLTLPNVRVGLPINQRFILVELSNHAELS 61
|||||

Db 181 TRIRYNRS 189
|||||
Db 181 TRIRYNRS 189

RESULT 4
US-10-083-336A-8
; Sequence 8, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnie, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 8
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-8

Query Match 98.5%; Score 956; DB 12; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.4e-101;
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTLPNVRVGLPINQRFILVELSNHAELS 63
Db 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTLPNVRVGLPINQRFILVELSNHAELS 62
Qy 64 ALDVTNAYVVGVRAGNSAYFFHFDNQEDAEAIHTLFTDVQNRVYTFAGGNYDRLEQL 123
Db 63 ALDVTNAYVVGVRAGNSAYFFHFDNQEDAEAIHTLFTDVQNRVYTFAGGNYDRLEQL 122
Qy 124 LRENIEGNGLPBEAISALYYSTGGTQPLTLARSFFIICIMISEAARFOYIEGEMR 183
Db 123 LRENIEGNGLPBEAISALYYSTGGTQPLTLARSFFIICIMISEAARFOYIEGEMR 182
Qy 184 RYNRS 189
Db 183 RYNRS 188

RESULT 5
US-10-083-336A-5
; Sequence 5, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrnie, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
```



```
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-5

Query Match      98.5%; Score 956; DB 12; Length 199;
Best Local Similarity 95.0%; Pred. No. 1.5e-101;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50
DB 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFIL 60
QY 51 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAF 110
DB 61 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAF 120
QY 111 GGNVDRLEQLAGNLRNENIELGNGLPLEEASALYYSTGGTQPLTLARSFIICQMISEAA 170
DB 121 GGNVDRLEQLAGNLRNENIELGNGLPLEEASALYYSTGGTQPLTLARSFIICQMISEAA 180
QY 171 RFQYIEGEMRTRIRYNRRS 189
DB 181 RFQYIEGEMRTRIRYNRRS 199

RESULT 6
US-10-083-336A-10
; Sequence 10, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-10

Query Match      98.5%; Score 956; DB 12; Length 200;
Best Local Similarity 95.0%; Pred. No. 1.5e-101;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFIL 50
DB 1 MIFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFIL 60
QY 51 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAF 110
DB 61 VELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAF 120
QY 111 GGNVDRLEQLAGNLRNENIELGNGLPLEEASALYYSTGGTQPLTLARSFIICQMISEAA 170
DB 121 GGNVDRLEQLAGNLRNENIELGNGLPLEEASALYYSTGGTQPLTLARSFIICQMISEAA 180
QY 171 RFQYIEGEMRTRIRYNRRS 189
DB 181 RFQYIEGEMRTRIRYNRRS 199

RESULT 7
US-10-083-336A-3
; Sequence 3, Application US/10083336A
; Publication No. US20030181665A1
; GENERAL INFORMATION:
; APPLICANT: Olson, Mark A
; APPLICANT: Millard, Charles B
; APPLICANT: Byrne, Michael P
; APPLICANT: Wannemacher, Robert W
; TITLE OF INVENTION: Ricin Vaccine and Methods of Making and Using Thereof
; FILE REFERENCE: P67452US0 (RIID 01-58)
; CURRENT APPLICATION NUMBER: US/10/083,336A
; CURRENT FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Ricinus communis
US-10-083-336A-3

Query Match      97.9%; Score 951; DB 12; Length 198;
Best Local Similarity 94.9%; Pred. No. 5.7e-101;
Matches 189; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINQRFILV 51
DB 1 IFPKQYPIINFTTAGATVQSYTNFIRAVRGRLTGGADVRHEIPVLPNRVGLPINQRFILV 60
QY 52 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFG 111
DB 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFG 120
QY 112 GNVDRLEQLAGNLRNENIELGNGLPLEEASALYYSTGGTQPLTLARSFIICQMISEAA 171
DB 121 GNVDRLEQLAGNLRNENIELGNGLPLEEASALYYSTGGTQPLTLARSFIICQMISEAA 180
QY 172 FOYIEGEMRTRIRYNRRS 189
DB 181 FOYIEGEMRTRIRYNRRS 198

RESULT 8
US-10-127-890-1
; Sequence 1, Application US/10127890
; Publication No. US20030166196A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; APPLICANT: Carroll, Stephen F.
; APPLICANT: Studnika, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127,890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/988,430
; FILING DATE: 09-DEC-1992
```

Query Match	97.9%;	Score 951;	DB 12;	Length 576;
Best Local Similarity	94.9%;	Pred. No. 2.6e-100;		
Matches 188;	Conservative 0;	Mismatches 0;	Indels 10;	Gaps 1;

Qy	2	IFPKQYPIINF	TGATGQSYTNF	IRAVGRLLT	-----	VLPNVRGLPINQRFILV	51
Db	36	IFPKQYPIINF	TGATGQSYTNF	IRAVGRLLT	TTGADVRRHEIPVLPNVRGLPINQRFILV	95	
Qy	52	ELSNHAEISVL	ALDVTHNAYV	VVGYGAGNSA	FFHHQDNCQDEA	EAITHLTDTQVNR	TFAG 111

RESULT 11
US-10-282-935-1
; Sequence 1, Application US/10282935
; Publication No. US20030143193A1
; GENERAL INFORMATION:
; APPLICANT: VITETTA, ELLEN S.
; APPLICANT: GHETTIE, VICTOR F.
; APPLICANT: SMALLSHAW, JOAN
; APPLICANT: BALUNA, ROXANA G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR MODIFYING TOXIC EFFECTS OF
; TITLE OF INVENTION: PROTEINACEOUS COMPOUNDS
; FILE REFERENCE: UTSD:884US
; CURRENT APPLICATION NUMBER: US/10/282,935
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 09/538,873
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/126,826
; PRIOR FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1

Search completed: February 10, 2004, 16:53:53
Job time : 25.3781 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:35 ; Search time 10.2276 Seconds
(without alignments)
1777.145 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971

Sequence: 1 MIFPKQYPIINFITAGATVQ.....ARFOYIEGEMTRIRYNERS 189

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: Pirl1.*

2: Pirl2.*

3: Pirl3.*

4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	951	97.9	576	1	RLCSD
2	861.5	88.7	564	1	ricin D precursor
3	336	34.6	289	1	agglutinin precurs
4	327.5	33.7	528	2	rRNA N-glycosidase
5	327.5	33.7	562	2	abrin-d precursor
6	326	33.6	247	2	abrin-c precursor
7	326	33.6	247	2	karasurin-B - Mongol
8	326	33.6	247	2	karasurin C - Tric
9	323	33.3	527	2	abrin-b precursor
10	310.5	32.0	251	2	abrin (clone 7.2)
11	303.5	31.3	528	1	abrin-a precursor
12	300.5	30.9	278	2	beta-luffin - smoo
13	293.5	30.2	250	2	luffin-b - smooth
14	274.5	28.3	277	2	rRNA N-glycosidase
15	273	28.1	286	1	rRNA N-glycosidase
16	273	28.1	570	2	agglutinin I precu
17	272	28.0	286	2	rRNA N-glycosidase
18	270	27.8	245	2	rRNA N-glycosidase
19	265	27.3	286	2	rRNA N-glycosidase
20	264	27.2	254	2	mistletoe lectin I
21	243.5	25.1	316	2	rRNA N-glycosidase
22	186.5	19.2	294	2	rRNA N-glycosidase
23	174	17.9	278	2	rRNA N-glycosidase
24	170	17.5	313	2	rRNA N-glycosidase
25	168.5	17.4	261	2	antiviral protein
26	148.5	15.3	289	2	rRNA N-glycosidase
27	138	14.2	272	2	betavulgin - beet
28	131	13.5	253	2	rRNA N-glycosidase
29	127	13.1	253	2	rRNA N-glycosidase

ALIGNMENTS

RESULT 1

RLCSD

ricin D precursor - castor bean

N:Contains: rRNA N-glycosidase (EC 3.2.2.22)

C:Species: Ricinus communis (castor bean)

C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999

C:Accession: A24041; S20513; A24614; A03372; A24010; A03374; S10903

R:Halling, K.C.; Halling, A.C.; Murray, E.E.; Ladin, B.F.; Houston, L.L.; Weaver, R.P.

Nucleic Acids Res. 13, 8019-8033, 1985

A:Title: Genomic cloning and characterization of a ricin gene from Ricinus communis.

A:Reference number: A24041; MUID:86067214; PMID:2959712

A:Accession: A24041

A:Molecule type: DNA

A:Residues: 1-576 <HAL>

A:Cross-references: GB:X03179; NID:g21082; PIDN:CAA26939.1; PID:g21083

R:Tregear, J.W.; Roberts, L.M.

Plant Mol. Biol. 18, 515-525, 1992

A:Title: The lectin gene family of Ricinus communis: cloning of a functional ricin gene

A:Reference number: S20513; MUID:92163016; PMID:1371405

A:Accession: S20513

A:Molecule type: DNA

A:Residues: 1-576 <TRE>

A:Cross-references: EMBL:X52908; NID:g21084; PIDN:CAA37095.1; PID:g21085

R:Lamb, F.I.; Roberts, L.M.; Lord, J.M.

Eur. J. Biochem. 148, 265-270, 1985

A:Title: Nucleotide sequence of cloned cDNA coding for preproricin.

A:Reference number: A24614; MUID:85179479; PMID:3838723

A:Accession: A24614

A:Molecule type: mRNA

A:Residues: 12-75, 'D', 77-550, 'R', 552-576 <LAM>

A:Cross-references: GB:X02388; NID:g21077; PIDN:CAA26230.1; PID:g21078

R:Yoshitake, S.; Funatsu, G.; Funatsu, M.

Agric. Biol. Chem. 42, 1267-1274, 1978

A:Title: Isolation and sequences of peptic peptides, and the complete sequence of Ile cha

A:Reference number: A03372

A:Accession: A03372

A:Molecule type: protein

A:Residues: 36-97, 'Q', 99-109, 'S', 111-269, 'D', 272-283, 'L', 285-288, 290-302 <YOS>

A:Note: this paper cites the others in the series providing experimental details for the

R:Araki, T.; Funatsu, G.

FEBS Lett. 191, 121-124, 1985

A:Title: Revised amino acid sequence of the B-chain of ricin D due to loss of tryptophan

A:Reference number: A24010

A:Accession: A24010

A:Molecule type: protein

A:Residues: 315-383, 'PS', 386-576 <ARA>

R:Funatsu, G.; Kimura, M.; Funatsu, M.

Agric. Biol. Chem. 43, 2221-2224, 1979

A:Title: Primary structure of Ala chain of ricin D.

A:Reference number: A03374

A:Accession: A03374

A:Molecule type: protein

R:Araki, T.; Yoshioka, Y.; Funatsu, G.
Biochim. Biophys. Acta 870, 277-285, 1986
A:Title: The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin
A:Reference number: A24210
A:Accession: A24210
A:Molecule type: protein
A:Residues: 303-325, 'F', 327-330, 'T', 332-361, 'D', 363-373, 'G', 375-403, 'T', 405-551, 'V', 553-564, 'A', 566-580, 'G', 582-599, 'D', 601-610, 'G', 612-621, 'G', 623-632, 'G', 634-643, 'G', 645-654, 'G', 656-665, 'G', 667-676, 'G', 678-687, 'G', 689-698, 'G', 700-709, 'G', 711-720, 'G', 722-731, 'G', 733-742, 'G', 744-753, 'G', 755-764, 'G', 766-775, 'G', 777-786, 'G', 788-797, 'G', 799-808, 'G', 810-819, 'G', 821-830, 'G', 832-841, 'G', 843-852, 'G', 854-863, 'G', 865-874, 'G', 876-885, 'G', 887-896, 'G', 898-907, 'G', 909-918, 'G', 920-929, 'G', 931-940, 'G', 942-951, 'G', 953-962, 'G', 964-973, 'G', 975-984, 'G', 986-995, 'G', 997-1006, 'G', 1008-1017, 'G', 1019-1028, 'G', 1030-1039, 'G', 1041-1050, 'G', 1052-1061, 'G', 1063-1072, 'G', 1074-1083, 'G', 1085-1094, 'G', 1096-1105, 'G', 1107-1116, 'G', 1118-1127, 'G', 1129-1138, 'G', 1140-1149, 'G', 1151-1160, 'G', 1162-1171, 'G', 1173-1182, 'G', 1184-1193, 'G', 1195-1204, 'G', 1206-1215, 'G', 1217-1226, 'G', 1228-1237, 'G', 1239-1248, 'G', 1250-1259, 'G', 1261-1270, 'G', 1272-1281, 'G', 1283-1292, 'G', 1294-1303, 'G', 1305-1314, 'G', 1316-1325, 'G', 1327-1336, 'G', 1338-1347, 'G', 1349-1358, 'G', 1360-1369, 'G', 1371-1380, 'G', 1382-1391, 'G', 1393-1402, 'G', 1404-1413, 'G', 1415-1424, 'G', 1426-1435, 'G', 1437-1446, 'G', 1448-1457, 'G', 1459-1468, 'G', 1470-1479, 'G', 1481-1490, 'G', 1492-1501, 'G', 1503-1512, 'G', 1514-1523, 'G', 1525-1534, 'G', 1536-1545, 'G', 1547-1556, 'G', 1558-1567, 'G', 1569-1578, 'G', 1580-1589, 'G', 1591-1600, 'G', 1602-1611, 'G', 1613-1622, 'G', 1624-1633, 'G', 1635-1644, 'G', 1646-1655, 'G', 1657-1666, 'G', 1668-1677, 'G', 1679-1688, 'G', 1690-1699, 'G', 1701-1710, 'G', 1712-1721, 'G', 1723-1732, 'G', 1734-1743, 'G', 1745-1754, 'G', 1756-1765, 'G', 1767-1776, 'G', 1778-1787, 'G', 1789-1798, 'G', 1800-1809, 'G', 1811-1820, 'G', 1822-1831, 'G', 1833-1842, 'G', 1844-1853, 'G', 1855-1864, 'G', 1866-1875, 'G', 1877-1886, 'G', 1888-1897, 'G', 1899-1908, 'G', 1910-1919, 'G', 1921-1930, 'G', 1932-1941, 'G', 1943-1952, 'G', 1954-1963, 'G', 1965-1974, 'G', 1976-1985, 'G', 1987-1996, 'G', 1998-2007, 'G', 2009-2018, 'G', 2020-2029, 'G', 2031-2040, 'G', 2042-2051, 'G', 2053-2062, 'G', 2064-2073, 'G', 2075-2084, 'G', 2086-2095, 'G', 2097-2106, 'G', 2108-2117, 'G', 2119-2128, 'G', 2130-2139, 'G', 2141-2150, 'G', 2152-2161, 'G', 2163-2172, 'G', 2174-2183, 'G', 2185-2194, 'G', 2196-2205, 'G', 2207-2216, 'G', 2218-2227, 'G', 2229-2238, 'G', 2240-2249, 'G', 2251-2260, 'G', 2262-2271, 'G', 2273-2282, 'G', 2284-2293, 'G', 2295-2304, 'G', 2306-2315, 'G', 2317-2326, 'G', 2328-2337, 'G', 2339-2348, 'G', 2350-2359, 'G', 2361-2370, 'G', 2372-2381, 'G', 2383-2392, 'G', 2394-2403, 'G', 2405-2414, 'G', 2416-2425, 'G', 2427-2436, 'G', 2438-2447, 'G', 2449-2458, 'G', 2460-2469, 'G', 2471-2480, 'G', 2482-2491, 'G', 2493-2502, 'G', 2504-2513, 'G', 2515-2524, 'G', 2526-2535, 'G', 2537-2546, 'G', 2548-2557, 'G', 2559-2568, 'G', 2570-2579, 'G', 2581-2590, 'G', 2592-2601, 'G', 2603-2612, 'G', 2614-2623, 'G', 2625-2634, 'G', 2636-2645, 'G', 2647-2656, 'G', 2658-2667, 'G', 2669-2678, 'G', 2680-2689, 'G', 2691-2700, 'G', 2702-2711, 'G', 2713-2722, 'G', 2724-2733, 'G', 2735-2744, 'G', 2746-2755, 'G', 2757-2766, 'G', 2768-2777, 'G', 2779-2788, 'G', 2790-2800, 'G', 2802-2811, 'G', 2813-2822, 'G', 2824-2833, 'G', 2835-2844, 'G', 2846-2855, 'G', 2857-2866, 'G', 2868-2877, 'G', 2879-2888, 'G', 2890-2899, 'G', 2901-2910, 'G', 2912-2921, 'G', 2923-2932, 'G', 2934-2943, 'G', 2945-2954, 'G', 2956-2965, 'G', 2967-2976, 'G', 2978-2987, 'G', 2989-2998, 'G', 3000-3009, 'G', 3011-3020, 'G', 3022-3031, 'G', 3033-3042, 'G', 3044-3053, 'G', 3055-3064, 'G', 3066-3075, 'G', 3077-3086, 'G', 3088-3097, 'G', 3099-3108, 'G', 3110-3119, 'G', 3121-3130, 'G', 3132-3141, 'G', 3143-3152, 'G', 3154-3163, 'G', 3165-3174, 'G', 3176-3185, 'G', 3187-3196, 'G', 3198-3207, 'G', 3209-3218, 'G', 3220-3229, 'G', 3231-3240, 'G', 3242-3251, 'G', 3253-3262, 'G', 3264-3273, 'G', 3275-3284, 'G', 3286-3295, 'G', 3297-3306, 'G', 3308-3317, 'G', 3319-3328, 'G', 3330-3339, 'G', 3341-3350, 'G', 3352-3361, 'G', 3363-3372, 'G', 3374-3383, 'G', 3385-3394, 'G', 3396-3405, 'G', 3407-3416, 'G', 3418-3427, 'G', 3429-3438, 'G', 3440-3449, 'G', 3451-3460, 'G', 3462-3471, 'G', 3473-3482, 'G', 3484-3493, 'G', 3495-3504, 'G', 3506-3515, 'G', 3517-3526, 'G', 3528-3537, 'G', 3539-3548, 'G', 3550-3559, 'G', 3561-3570, 'G', 3572-3581, 'G', 3583-3592, 'G', 3594-3603, 'G', 3605-3614, 'G', 3616-3625, 'G', 3627-3636, 'G', 3638-3647, 'G', 3649-3658, 'G', 3660-3669, 'G', 3671-3680, 'G', 3682-3691, 'G', 3693-3702, 'G

[illegible]

QY 182 RI 183
|:
Db 174 RV 175

RESULT 8
JCS606
karasurin C - Trichosanthes kirilowii var. japonica
N;Contains: Karasurin A
C;Species: Trichosanthes kirilowii var. japonica
C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 19-Jul-2002
C;Accession: JCS606; JC5033
R;Mizukami, H.; Iida, K.; Kondo, T.; Ogihara, Y.
BioI. Pharm. Bull. 20, 711-713, 1997
A;Title: Cloning and bacterial expression of a gene encoding ribosome-inactivating prote
A;Reference number: JCS606; MUID:97356562; PMID:9212998
A;Accession: JCS606

A;Molecule type: DNA
A;Residues: 1-289 <MAX>
A;Cross-references: DDBJ:AB000666; NID:g2329830; PIDN:BAA21786.1; PID:g2329831
R;Kondo, T.; Mizukami, H.; Takeda, T.; Ogihara, Y.
Biol. Pharm. Bull. 19, 1485-1489, 1996
A;Title: Amino acid sequences and ribosome-inactivating activities of karasurin-B and ka
A;Reference number: JCS032; MUID:97109848; PMID:8951169
A;Accession: JCS033
A;Status: preliminary
A;Molecule type: protein
A;Residues: 22-270 <KON>
C;Comment: This protein is a ribosome-inactivating protein and exhibit cytotoxic, aborti
C;Comment: This protein belongs to type I ribosomal-inactivating proteins which catalyti
C;Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
F;22-270/Product: karasurin C #status predicted <MAC>
F;24-270/Product: karasurin A #status predicted <MAC>
F;27-266/Domain: rRNA N-glycosidase homology <RNG>

Query Match 33.6%; Score 326; DB 2; Length 289;
Best Local Similarity 39.6%; Pred. No. 2.3e-22;
Matches 72; Conservative 44; Mismatches 50; Indels 16; Gaps 5;

[illegible]

A:Title: The complete amino acid sequence of the A-chain of abrin-a, a toxic protein from
A:Reference number: J0202
A:Accession: J0202
A:Molecule type: protein
A:Residues: 1-201,203-251 <FUN>
A:Note: the amino-terminal residue forms pyrrolidone carboxylic acid; therefore, we have
R:Evans, G.; Mathiesen, A.; Sundan, A.
J. Biol. Chem. 266, 6848-6852, 1991
A:Title: Direct molecular cloning and expression of two distinct abrin A-chains.
A:Reference number: A39761; MUID:91201329; PMID:2016300
A:Accession: A39761
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 'E',2-251 <EVE>
A:Cross-references: GB:54872
A:Note: residues 1-8 were derived from the synthesized primer
R:Kimura, M.; Sumizawa, T.; Funatsu, G.
Biosci. Biotechnol. Biochem. 57, 166-169, 1993
A:Title: The complete amino acid sequences of the B-chains of Abrin-a and Abrin-b, toxic
A:Reference number: J01398; MUID:93169023; PMID:7763422
A:Contents: seeds
A:Accession: J01398
A:Molecule type: protein
A:Residues: 261-347, 'T',349-351, 'A',353-357, 'L',359-528 <KIM>
A:Experimental source: seed
R:Evans, G.; Mathiesen, A.; Sundan, A.
submitted to the EMBL Data Library, October 1990
A:Description: Direct molecular cloning of two distinct abrin A-chains.
A:Reference number: S14471
A:Accession: S14472
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 'ME',2-251 <EV2>
A:Cross-references: EMBL:X54873; NID:g16090; PIDN:CAA38655.1; PID:g16091
R:Chen, Y.L.; Chow, L.P.; Tsugita, A.; Lin, J.Y.
FEBS Lett. 309, 115-118, 1992
A:Title: The complete primary structure of abrin-a B chain.
A:Reference number: S24133; MUID:92371656; PMID:1505674
A:Accession: S24133
A:Molecule type: protein
A:Residues: 262-297, 'Y',299-426, 'L',428-466, 'P',468-482, 'L',484-528 <CHE>
R:Lin, S.H.; Chow, L.P.; Chen, Y.L.; Liaw, Y.C.; Chen, J.K.; Lin, J.Y.
Eur. J. Biochem. 240, 564-569, 1996
A:Title: Probing the domain structure of abrin-a by tryptic digestion.
A:Reference number: S74110; MUID:97008945; PMID:8856055
A:Accession: S74110
A:Molecule type: protein
A:Residues: 89-108;154-172 <LIN>
A:Experimental source: seed
A:Accession: S74111
A:Molecule type: protein
A:Residues: 262-276, 'X',278-280;329-348;369-388;399-418 <LIW>
A:Experimental source: seed
C:Comment: Abrin-a is more toxic than ricin. The toxin consists of an A chain, which inhibits
taining receptors on the cell surface. The A and B chains are linked by a single disulfide
C:Superfamily: ricin; rRNA N-glycosidase homology
C:Keywords: duplication; glycoprotein; glycosidase; hydrolase; lectin; pyroglutamic acid
F:1-251/Product: abrin-a chain A #status experimental <ACH>
F:7-245/Domain: rRNA N-glycosidase homology <RNG>
F:261-528/Product: abrin-a chain B #status experimental <BCH>
F:283-325;326-366;369-407;414-449;453-492;495-528/Region: 40-residue repeats
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F:74,113,195,196/Binding site: substrate (Tyr, Tyr, Glu, Asn) #status predicted
F:164,167/Active site: Glu, Arg #status predicted
F:247-269,286-305,329-346,417-430,456-473/Disulfide bonds: #status predicted
F:288,312/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted
F:361,401/Binding site: carboxylate (Asn) (covalent) #status experimental
F:500,521/Binding site: N-acetylglucosamine (Asp, Asn) #status predicted

Query Match 31.3%; Score 303.5; DB 1; Length 528;
Best Local Similarity 41.1%; Pred. No. 5.7e-20;
Matches 76; Conservative 26; Mismatches 64; Indels 19; Gaps 5;

QY 10 INFETAGATVQSYTNFIRAVGRGL-----TVLPNRVGLPINQRFILVLSNHAELSV 61
DB 5 IKFSTEGATQSYKQFIETALRERLGGLIHDPVLPDPTTQERNRYITVLSNSDTSI 64
QY 62 TLALDVTNAYVVGVRAGNSAYFFH--PDNQDAEAIHLFTDVQNRVYFAFGGNYDRLEQ 119
DB 65 EVGIDVTNAYVAVRAGTQSYFLRDAESSASD-----YLFGT-DQHSLPFGTYGDLER 118
QY 120 LAGNIRENIEGNGLPEBAISALYYSTGGTQLPFLARSFFICIQMISEARFQVIEGEM 179
DB 119 WAHQSQOIPGLQALTHGIS---FFRSGGNDNEKARTLIVIQMAEAARFRYISNRV 175
QY 180 RTRIR 184
DB 176 RVSQ 180

RESULT 12

S23519
beta-luffin - smooth loofah
C:Species: Luffa cylindrica (smooth loofah)
C:Date: 17-Apr-1993 #sequence revision 17-Apr-1993 #text_change 20-Aug-1999
C:Accession: S23519; S23113
R:Kataoka, J.; Habuka, N.; Miyano, M.; Masuta, C.; Koizumi, A.
Plant Mol. Biol. 19, 887-889, 1992
A:Title: Nucleotide sequence of cDNA encoding beta-luffin, another ribosome-inactivating
A:Reference number: S23519; MUID:92353400; PMID:1643290
A:Accession: S23519
A:Molecule type: mRNA
A:Residues: 1-278 <KAT>
A:Cross-references: EMBL:X62372; NID:g19149; PIDN:CAA44230.1; PID:g19150
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
F:26-264/Domain: rRNA N-glycosidase homology <RNG>

Query Match 30.9%; Score 300.5; DB 2; Length 278;
Best Local Similarity 35.3%; Pred. No. 4.7e-20;
Matches 66; Conservative 43; Mismatches 21; Gaps 4;

QY 10 INFETAGATVQSYTNFIRAVGRGL-----TVLPNRVGLPINQRFILVLSNHAEL 59
DB 24 VSFSLGADSKSYKFTALRKALPSKSVNIPLLPSASGA---SRVILMQLSNYDAK 80
QY 60 SVTLALDVTNAYVVGVRAGNSAYFFHFDNQDAEAIHLFTDVQNRVYFAFGGNYDRLEQ 119
DB 81 AITMAIDVTNVIMGYLVNSTSYFF--NESDAKLASQVYFKGSTIVLPYSGNYERLQN 137
QY 120 LAGNIRENIEGNGLPEBAISALYYSTGGTQLPFLARSFFICIQMISEARFQVIEGEM 179
DB 138 AAGKVREKIPGLGFRAFDSAITSLPHYDS-----TAAGAFILVIQTABASRFKIEGQI 192
QY 180 RTRIRVN 186
DB 193 IERIPKN 199

RESULT 13

JN0108
luffin-b - smooth loofah
C:Species: Luffa cylindrica (smooth loofah)
C:Date: 04-Sep-1998 #sequence revision 04-Sep-1998 #text_change 07-May-1999
C:Accession: JN0108
R:Islam, M.R.; Hirayama, H.; Funatsu, G.
Agric. Biol. Chem. 55, 229-238, 1991
A:Title: Complete amino acid sequence of luffin-b, a ribosome-inactivating protein from
A:Reference number: JN0108; MUID:91248488; PMID:1368666
A:Accession: JN0108
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-250 <ISL>
C:Superfamily: rRNA N-glycosidase; rRNA N-glycosidase homology
F:5-246/Domain: rRNA N-glycosidase homology <RNG>

Query Match 30.2%; Score 293.5; DB 2; Length 250;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:55 ; Search time 6.30467 Seconds
(without alignments)
1409.756 Million cell updates/sec

Title: US-10-083-336A-6
Perfect score: 971
Sequence: 1 MIFPKQYPLINFTTAGATVQ.....ARQYIEGEMRTIRYNRS 189

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	951	97.9	576	1 RICI_RICCO	P02879 ricinus com
2	861.5	88.7	564	1 AGGL_RICCO	P06750 ricinus com
3	336	34.6	289	1 RIPC_TRIKI	P09989 trichosanthe
4	327.5	33.7	562	1 ABRC_ABRPR	P28590 abrus prece
5	326	33.6	289	1 RIPS_TRIKI	P24478 trichosanthe
6	323	33.3	527	1 ABRB_ABRPR	Q66077 abrus prece
7	320.5	33.0	282	1 RIPC_BRVDI	P98184 bryonia dio
8	307.5	31.7	563	1 NIGB_SAMNI	P33183 sambucus ni
9	305	31.4	290	1 RIPC_BRVDI	P33185 bryonia dio
10	303.5	31.3	528	1 ABRA_ABRPR	P11140 abrus prece
11	293.5	30.2	250	1 RIPC_LUCFY	P22851 luffa cylin
12	281.5	29.0	286	1 RIPC_CUCFI	Q9frx4 cucumis fig
13	274.5	28.3	277	1 RIPC_LUCFY	Q00465 luffa cylin
14	273	28.1	286	1 RIPC_MONCH	P16094 momordica c
15	272	28.0	286	1 RIPC_MONCH	P29339 momordica b
16	264	27.2	254	1 MIA_VISAL	P81446 viscum albu
17	257	26.5	294	1 RIPC_TRIAN	P56626 trichosanthe
18	243.5	25.1	316	1 RIPC_GELMU	P33186 gelonium mu
19	186.5	19.2	294	1 RIPC_PHYAM	Q03464 phytoacca
20	174	17.9	278	1 RIPC_MIRUA	P21326 mirabilis j
21	170	17.5	313	1 RIPC_PHYAM	P10297 phytoacca
22	168.5	17.4	261	1 RIPC_PHYAM	P23339 phytoacca
23	131	13.5	253	1 RIPC_SAPOF	Q41391 saponaria o
24	127	13.1	253	1 RIPC_SAPOF	Q41389 saponaria o
25	125	12.9	310	1 RIPC_PHYAM	Q40772 phytoacca
26	124	12.8	299	1 RIPC_SAPOF	P20656 saponaria o
27	121	12.5	292	1 RIPC_SAPOF	P27559 saponaria o
28	116.5	12.0	293	1 RIPC_DIACA	P24476 dianthus ca
29	114	11.7	280	1 RIPC_HORVU	P04399 hordeum vul
30	110	11.3	236	1 RIPC_SAPOF	P27560 saponaria o
31	108	11.1	319	1 SLTA_BP933	P09385 bacterioph
32	107	11.0	280	1 RIPC_HORVU	P22244 hordeum vul
33	103.5	10.7	300	1 RIPC_MAIZE	P25891 zea mays (m

RESULT 1

RICI_RICCO STANDARD; PRT; 576 AA.
 AC P02879; P02880;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ricin precursor [Contains: Ricin A chain (rRNA N-glycosidase)
 (EC 3.2.2.22); Ricin B chain].
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
 CX NCBI_TaxID=3988;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=86067214; PubMed=2999712;
 RA Halling K.C., Halling A.C., Murray E.E., Ladin B.F., Houston L.L.,
 RA Weaver R.F.;
 RT "Genomic cloning and characterization of a ricin gene from Ricinus
 communis";
 RL Nucleic Acids Res. 13:8019-8033(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92163016; PubMed=1371405;
 RA Tregear J.W., Roberts L.M.;
 RT "The lectin gene family of Ricinus communis: cloning of a functional
 ricin gene and three lectin pseudogenes.";
 RL Plant Mol. Biol. 18:515-525(1992).
 RN [3]
 RP SEQUENCE OF 12-576 FROM N.A.
 RX MEDLINE=85179479; PubMed=3838723;
 RA Lamb A., Roberts L.M., Lord J.M.;
 RT "Nucleotide sequence of cloned cDNA coding for preproricin.";
 RL Eur. J. Biochem. 148:265-270(1985).
 RN [4]
 RP SEQUENCE OF 36-302.
 RX Yoshitake S., Funatsu G., Funatsu M.;
 RT "Isolation and sequences of peptic peptides, and the complete
 sequence of ile chain of ricin-D.";
 RL Agric. Biol. Chem. 42:1267-1274(1978).
 RN [5]
 RP SEQUENCE OF 315-576.
 RX Funatsu G., Kimura M., Funatsu M.;
 RT "Primary structure of Ala chain of ricin D.";
 RL Agric. Biol. Chem. 43:2221-2224(1979).
 RN [6]
 RP CARBOHYDRATE-LINKAGE SITES, AND PARTIAL SEQUENCE.
 RX MEDLINE=90344223; PubMed=1368517;
 RA Kimura Y., Kusuoku H., Tada M., Takagi S., Funatsu G.;
 RT "Structural analyses of sugar chains from ricin A-chain variant.";
 RL Agric. Biol. Chem. 54:157-162(1990).
 RN [7]
 RP REVIEW.
 RX MEDLINE=21480122; PubMed=11595634;
 RA Olsnes S., Kozlov J.V.;

ALIGNMENTS

RT "Ricin."; CC
 RL Toxicon 39:1723-1728 (2001). CC
 RN [8] CC
 RX X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS). CC
 RP MEDLINE=87165983; PubMed=3558397; CC
 RA Monfort W., Villafraña J.E., Monzingo A.F., Ernst S.R., Katzin B., CC
 RA Rutenber E., Xuong N.H., Hamlin R., Robertus J.D.; CC
 RA "The three-dimensional structure of ricin at 2.8 Å."; CC
 RL J. Biol. Chem. 262:5398-5403 (1987). CC
 RN [9] CC
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN. CC
 RX MEDLINE=91352004; PubMed=1881881; CC
 RA Katzin B.J., Collins E.J., Robertus J.D.; CC
 RA "Structure of ricin A-chain at 2.5 Å."; CC
 RL Proteins 10:251-259 (1991). CC
 RN [10] CC
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF B-CHAIN. CC
 RX MEDLINE=91352005; PubMed=1881882; CC
 RA Rutenber E., Robertus J.D.; CC
 RA "Structure of ricin B-chain at 2.5-Å resolution."; CC
 RL Proteins 10:260-269 (1991). CC
 RN [11] CC
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF A-CHAIN. CC
 RX MEDLINE=95082010; PubMed=7990130; CC
 RA Weston S.A., Tucker A.D., Thatcher D.R., Derbyshire D.J., CC
 RA Paucit R.A.; CC
 RA "X-ray structure of recombinant ricin A-chain at 1.8-Å resolution."; CC
 RL J. Mol. Biol. 244:410-422 (1994). CC
 RN [12] CC
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT HIS-215. CC
 RX MEDLINE=96374222; PubMed=8780513; CC
 RA Day P.J., Ernst S.R., Frankel A.E., Monzingo A.F., Pascal J.M., CC
 RA Molina-Svinth M.C., Robertus J.D.; CC
 RA "Structure and activity of an active site substitution of ricin A CC
 RT chain."; CC
 RL Biochemistry 35:11098-11103 (1996). CC
 RN [13] CC
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF A-CHAIN. CC
 RX MEDLINE=97240820; PubMed=9086280; CC
 RA Yan X., Hollis T., Svinth M., Day P., Monzingo A.F., Milne G.W., CC
 RA Robertus J.D.; CC
 RA "Structure-based identification of a ricin inhibitor."; CC
 RL J. Mol. Biol. 266:1043-1049 (1997). CC
 RN [14] CC
 RP MUTAGENESIS. CC
 RX MEDLINE=93165632; PubMed=1287657; CC
 RA Xin Y., Robertus J.D.; CC
 RA "Analysis of several key active site residues of ricin A chain by CC
 RT mutagenesis and X-ray crystallography."; CC
 RL Protein Eng. 5:775-779 (1992). CC
 CC -!- FUNCTION: Ricin is highly toxic to animal cells and to a less CC
 CC extent to plant cells. The A chain is responsible for inhibiting CC
 CC protein synthesis through the catalytic inactivation of 60S CC
 CC ribosomal subunits. It acts as a glycosylase that removes a CC
 CC specific adenine residue from an exposed loop of 28S ribosomal CC
 CC RNA. As this loop is involved in the binding of elongation CC
 CC factors, the modified ribosomes are unable to support protein CC
 CC synthesis. The A chain can inactivate a few thousand ribosomes CC
 CC per minute, thus inactivating them faster than the cell can make CC
 CC new ones. A single A-chain molecule can therefore kill an animal CC
 CC cell. The B chain binds to cell receptors and facilitates the CC
 CC entry into the cell of the A chain; B chains are also responsible CC
 CC for cell agglutination (lectin activity). It binds to beta-D- CC
 CC galactopyranoside moieties. CC
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one CC
 CC specific adenosine on the 28S rRNA. CC
 CC -!- SUBUNIT: Disulfide-linked dimer of A and B chains. CC
 CC consists of 3 homologous subdomains (alpha, beta, gamma). CC
 CC -!- DOMAIN: The B chain is composed of two domains, each domain CC
 CC consists of 3 homologous subdomains (alpha, beta, gamma). CC
 CC -!- PTM: THE MAJOR A-CHAIN IS GLYCOSYLATED ONLY IN POSITION 45 AND THE CC
 CC MINOR A-CHAIN IS GLYCOSYLATED ALSO IN POSITION 271. CC
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME- CC
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY. CC

CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains. CC
 CC -!- CAUTION: REF.4 AND REF.5 SEQUENCES HAVE A HIGH NUMBER OF CONFLICTS CC
 CC WITH THE SEQUENCE TRANSLATED FROM DNA (REF.1, REF.2 AND REF.3). CC
 CC -!- DATABASE: NAME=Protein Spotlight; CC
 CC NOTE=Issue 31 of February 2003; CC
 CC WWW="http://www.expasy.org/spotlight/articles/sptlt031.html". CC
 CC ----- CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration CC
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation - CC
 CC the European Bioinformatics Institute. There are no restrictions on its CC
 CC use by non-profit institutions as long as its content is in no way CC
 CC modified and this statement is not removed. Usage by and for commercial CC
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ CC
 CC or send an email to license@isb-sib.ch). CC
 CC ----- CC
 CC EMBL; X03179; CAA26939.1; -; CC
 CC EMBL; X52908; CAA37095.1; -; CC
 CC EMBL; X02388; CAA26230.1; -; CC
 CC EMBL; A12892; CAA01058.1; -; CC
 CC EIR; A24041; RLCSD. CC
 CC PDB; 2AAI; 31-JAN-94. CC
 CC PDB; 1APG; 31-JAN-94. CC
 CC PDB; 1FMP; 31-OCT-93. CC
 CC PDB; 11FS; 14-JAN-98. CC
 CC PDB; 11FT; 14-JAN-98. CC
 CC PDB; 11FU; 14-JAN-98. CC
 CC PDB; 1R7C; 31-OCT-93. CC
 CC PDB; 1OBS; 16-JUN-97. CC
 CC PDB; 1OBT; 16-JUN-97. CC
 CC PDB; 1BR5; 02-SEP-98. CC
 CC PDB; 1BR6; 02-SEP-98. CC
 CC PDB; 1IL3; 16-JAN-02. CC
 CC PDB; 1IL4; 16-JAN-02. CC
 CC PDB; 1IU9; 16-JAN-02. CC
 CC GlycoSuiteDB; P02879; -; CC
 CC InterPro; IPR000772; Ricin_B_lectin. CC
 CC InterPro; IPR001574; RIP. CC
 CC Pfam; PF00652; Ricin_B_lectin; 6. CC
 CC Pfam; PF00161; RIP; 1. CC
 CC PRINTS; PRO0396; SHIGARICIN. CC
 CC SMART; SM00458; RICIN; 2. CC
 CC PROSITE; PS0231; RICIN_B_LECTIN; 2. CC
 CC PROSITE; PS0275; SHIGA_RICIN; 1. CC
 CC Plant defense: Hydrolase; Protein synthesis inhibitor; Toxin; Repeat; CC
 CC Glycoprotein; Lectin; Signal; 3D-structure. CC
 CC SIGNAL 1 35 CC
 CC CHAIN 36 302 CC
 CC PEPTIDE 303 314 CC
 CC CHAIN 315 576 CC
 CC CHAIN 321 448 CC
 CC DOMAIN 451 575 CC
 CC REPEAT 331 373 CC
 CC REPEAT 374 414 CC
 CC REPEAT 417 449 CC
 CC REPEAT 462 497 CC
 CC REPEAT 501 540 CC
 CC REPEAT 543 570 CC
 CC ACT_SITE 212 212 CC
 CC DISULFID 294 318 CC
 CC DISULFID 334 353 CC
 CC DISULFID 377 394 CC
 CC DISULFID 465 478 CC
 CC DISULFID 504 521 CC
 CC CARBOHYD 45 45 CC
 CC CARBOHYD 271 271 CC
 CC CARBOHYD 409 409 CC
 CC CARBOHYD 449 449 CC
 CC CARBOHYD 76 76 CC
 CC CONFLICT 551 551 CC
 CC STRAND 43 47 CC
 CC TURN 49 50 CC
 CC N-LINKED (GLCNAC. . .). CC
 CC /FTIG-CAR 00080. CC
 CC N-LINKED (GLCNAC. . .) (IN MINOR FORM). CC
 CC /FTIG-CAR 00081. CC
 CC N-LINKED (GLCNAC. . .). CC
 CC N-LINKED (GLCNAC. . .). CC
 CC E -> D (IN REF. 3). CC
 CC A -> R (IN REF. 3). CC

Query Match 97.9%; Score 951; DB 1; Length 576;
 Best Local Similarity 94.9%; Pred. No. 3.2e-80;
 Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY 2 IFPKQYPIINFNTAGATVQSYNFIKAVRGRLT-----VLPNRVGLPINQRFILV 51
 Db 36 IFPKQYPIINFNTAGATVQSYNFIKAVRGRLTGDVHRHEIPVLPNRVGLPINQRFILV 95

QY 52 ELSNHAELSVTLALDVTNAYVVGYNAGNSAYFFHPDNQDEAETHLFTDVQNRVYTFAG 111
 Db 96 ELSNHAELSVTLALDVTNAYVVGYNAGNSAYFFHPDNQDEAETHLFTDVQNRVYTFAG 155

QY 112 GNYDRLEQLAGNLENIENGPLEEASALYYSTGGTOLPTLARSFICQMISEAR 171
 Db 156 GNYDRLEQLAGNLENIENGPLEEASALYYSTGGTOLPTLARSFICQMISEAR 215

QY 172 FOYIEGEMTRIRYNRRS 189
 Db 216 FOYIEGEMTRIRYNRRS 233

RESULT 2
 AGGL RICCO STANDARD; PRT; 564 AA.
 AC P06750;
 DT 01-JAN-1988 (Rel. 06, Created)
 DT 01-JAN-1988 (Rel. 06, Last sequence update)
 DE Agglutinin precursor (RCA) [Contains: Agglutinin A chain (rRNA N-glycosidase) (BC 3.2.2.22); Agglutinin B chain].
 OS Ricinus communis (Castor bean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
 OX NCBI_TaxID=3988;
 RN [1]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=86059449; PubMed=2999130;
 RA Roberts L.M., Lamb F.T., Pappin D.J.C., Lord J.M.;
 RT "The primary sequence of Ricinus communis agglutinin. Comparison with ricin.";
 RN J. Biol. Chem. 260:15682-15686(1985).
 RN [2]
 RN SEQUENCE OF 303-564.
 RC TISSUE=Seed;
 RA Araki T., Yoshitaka Y., Funatsu G.;
 RT "The complete amino acid sequence of the B-chain of the Ricinus communis agglutinin isolated from large-grain castor bean seeds.";
 RN Biochim. Biophys. Acta 872:277-285(1986).
 RN [3]
 RN SEQUENCE OF 303-337.
 RX MEDLINE=80178723; PubMed=6768555;
 RA Lin T.-S., Li S.-L.;
 RT "Purification and physicochemical properties of ricins and agglutinins from Ricinus communis.";
 RN Eur. J. Biochem. 105:453-459(1980).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -1- SIMILARITY: Contains 2 ricin B-type lectin domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; M2089; AAA33869.1; -;
 CC EMBL; S40368; AAB22584.1; -;

DR PIR; A24261; RLCSAG.
 DR HSP; P02879; IBR6.
 DR GlycoSuiteDB; P06750; -;
 DR InterPro; IPR000772; Ricin_B_lectin.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00652; Ricin_B_lectin; 6.
 DR Pfam; PF00161; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS00231; RICIN_B_LECTIN; 2.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 KW plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 KW Glycoprotein; Lectin; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 290 AGGLUTININ A CHAIN.
 FT PROPEP 291 302 LINKER PEPTIDE.
 FT CHAIN 303 564 AGGLUTININ B CHAIN.
 FT DOMAIN 309 436 RICIN B-TYPE LECTIN 1.
 FT DOMAIN 439 563 RICIN B-TYPE LECTIN 2.
 FT REPEAT 319 361 1-ALPHA.
 FT REPEAT 362 402 1-BETA.
 FT REPEAT 405 437 1-GAMMA.
 FT REPEAT 450 485 2-ALPHA.
 FT REPEAT 489 528 2-BETA.
 FT REPEAT 531 558 2-GAMMA.
 FT ACT_SITE 200 200 BY SIMILARITY.
 FT DISULFID 282 306 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 322 341 BY SIMILARITY.
 FT DISULFID 365 382 BY SIMILARITY.
 FT DISULFID 453 466 BY SIMILARITY.
 FT DISULFID 492 509 BY SIMILARITY.
 FT CARBOHYD 34 34 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. .).
 FT CARBOHYD 437 437 N-LINKED (GLCNAC. .).
 FT CONFLICT 331 331 F -> T (IN REF. 2).
 FT CONFLICT 362 362 N -> D (IN REF. 2).
 FT CONFLICT 374 374 R -> G (IN REF. 2).
 FT CONFLICT 404 404 R -> T (IN REF. 2).
 FT CONFLICT 552 552 F -> V (IN REF. 2).
 SQ SEQUENCE 564 AA; 62851 MW; D455F2A72F609759 CRC64;

Query Match 88.7%; Score 861.5; DB 1; Length 564;
 Best Local Similarity 86.9%; Pred. No. 5.6e-72;
 Matches 172; Conservative 7; Mismatches 8; Indels 11; Gaps 2;

QY 2 IFPKQYPIINFNTAGATVQSYNFIKAVRGRLT-----VLPNRVGLPINQRFILV 51
 Db 25 IFPKQYPIINFNTAGATVQSYNFIKAVRGRLTGDVHRHEIPVLPNRVGLPINQRFILV 84

QY 52 ELSNHAELSVTLALDVTNAYVVGYNAGNSAYFFHPDNQDEAETHLFTDVQNRVYTFAG 111
 Db 85 ELSNHAELSVTLALDVTNAYVVGYNAGNSAYFFHPDNQDEAETHLFTDVQNRVYTFAG 144

QY 112 GNYDRLEQLAGNLENIENGPLEEASALYYSTGGTOLPTLARSFICQMISEAR 171
 Db 145 GNYDRLEQLAGNLENIENGPLEEASALYYSTGGTOLPTLARSFICQMISEAR 203

QY 172 FOYIEGEMTRIRYNRRS 189
 Db 204 FOYIEGEMTRIRYNRRS 221

RESULT 3
 RIPT TRIKI STANDARD; PRT; 289 AA.
 ID RIPT TRIKI
 AC P09889;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ribosome-inactivating protein alpha-trichosanthin precursor (rRNA N-glycosidase) (BC 3.2.2.22) (Alpha-TCS).
 OS Trichosanthes kirilowii (Mongolian snake-gourd).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales, Cucurbitaceae; Trichosanthes.
 ON NCBI_TaxID=3677;
 RX SEQUENCE FROM N.A.
 RC STRAIN=Maximowicz;
 RX MEDLINE=91153657; PubMed=199291;
 RA Shaw P.C., Yung M.H., Zhu R.H., Ho W.K.K., Ng T.B., Yeung H.W.;
 RT "Cloning of trichosanthin cDNA and its expression in *Escherichia*
 RL *coli*."; Gene 97:267-272(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Maximowicz; TISSUE=Leaf;
 RX MEDLINE=90256790; PubMed=2341400;
 RA Chow T., Feldman R.A., Lovett M., Piatak M.;
 RT "Isolation and DNA sequence of a gene encoding alpha-trichosanthin, a
 RL type I ribosome-inactivating protein."; J. Biol. Chem. 265:8670-8674(1990).
 RN [3]
 RP SEQUENCE OF 24-270.
 RC STRAIN=Maximowicz; TISSUE=Tuberous root;
 RX MEDLINE=90256789; PubMed=2341399;
 RA Collins E.J., Robertus J.D., Lopresti M., Stone K.L., Williams K.R.,
 Wu P., Hwang K., Piatak M.;
 RT "Primary amino acid sequence of alpha-trichosanthin and molecular
 RL models for abrin A-chain and alpha-trichosanthin."; J. Biol. Chem. 265:8665-8669(1990).
 RN [4]
 RP SEQUENCE OF 24-270.
 RC TISSUE=Tuberous root;
 RA Wang Y., Qian R.Q., Gu Z.W., Jin S.W., Zhang L.Q., Xia Z.X.,
 Tian G.Y., Ni C.Z.;
 RT "Scientific evaluation of Tian Hua Fen (THF): history, chemistry and
 RL application."; Pure Appl. Chem. 58:789-798(1986).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.88 ANGSTROMS).
 RX MEDLINE=94344957; PubMed=8066085;
 RA Zhou F., Fu Z., Chen M., Lin Y., Pan K.;
 RT "Structure of trichosanthin at 1.88-A resolution."; Proteins 19:4-13(1994).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
 RX MEDLINE=95344383; PubMed=7619070;
 RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
 RT "Studies on crystal structures, active-centre geometry and
 RL depurinatin mechanism of two ribosome-inactivating proteins."; Biochem. J. 309:285-298(1995).
 CC -!- FUNCTION: TRICHOSANTHIN IS AN ABORTION-INDUCING PROTEIN. IT IS
 CC CAPABLE OF INHIBITING HIV-1 INFECTION AND REPLICATION. IT
 CC INACTIVATES EUKARYOTIC 60S RIBOSOMAL SUBUNITS.
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 CC TYPE 1 RIP SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M34858; AAA34207.1; -;
 CC EMBL; J05434; AAA34206.1; -;
 CC PIR; J05666; RLTTZ.
 CC PDB; 1MRJ; 07-FEB-95.
 CC PDB; 1MRK; 07-FEB-95.
 CC PDB; 1TCS; 10-JUL-95.
 CC PDB; 1U4G; 28-JAN-03.

DR PDB; 1MLI; 21-JAN-03.
 DR PDB; 1QD2; 24-APR-00.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00161; RIP; 1.
 DR PRINTS; PRO0396; SHIGARICIN.
 DR PROSITE; PS00275; SHIGA_RICIN; 1.
 KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
 KW Toxin; Signal; 3D-structure.
 FT SIGNAL 1 23
 FT CHAIN 24 270
 FT PROPEP 271 289
 FT ACT_SITE 183
 FT CONFLICT 57 60
 FT CONFLICT 82 84
 FT CONFLICT 87 87
 FT CONFLICT 92 92
 FT CONFLICT 143 144
 FT CONFLICT 196 196
 FT CONFLICT 215 216
 FT CONFLICT 231 231
 FT CONFLICT 234 234
 FT CONFLICT 246 266
 FT CONFLICT 247 247
 FT STRAND 25 28
 FT TURN 30 31
 FT HELIX 34 46
 FT TURN 47 47
 FT STRAND 50 54
 FT TURN 55 56
 FT STRAND 57 60
 FT HELIX 66 69
 FT STRAND 70 76
 FT TURN 78 79
 FT STRAND 82 88
 FT TURN 89 92
 FT STRAND 93 99
 FT TURN 100 101
 FT STRAND 102 105
 FT HELIX 109 114
 FT TURN 115 117
 FT TURN 120 121
 FT STRAND 124 127
 FT HELIX 134 141
 FT TURN 142 142
 FT HELIX 145 147
 FT STRAND 150 150
 FT HELIX 152 163
 FT TURN 164 165
 FT HELIX 167 180
 FT TURN 181 181
 FT HELIX 182 186
 FT STRAND 187 187
 FT HELIX 188 195
 FT TURN 196 196
 FT STRAND 202 202
 FT HELIX 206 226
 FT TURN 227 230
 FT STRAND 231 239
 FT TURN 241 242
 FT STRAND 245 250
 FT TURN 251 252
 FT HELIX 254 258
 FT TURN 259 259
 FT STRAND 260 260
 FT STRAND 263 263
 FT TURN 266 268
 FT TURN 289 AA; 31676 MW; 5CE09BB630575BB9 CRC64;
 SQ SEQUENCE

Query Match 34.6%; Score 336; DB 1; Length 289;
 Best Local Similarity 38.9%; Pred. No. 7.9e-24;
 Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;

QY 10 INFTTAGATVQSYTNFIRAVRGLTVLPN-----RVGUPINQRFILVELSNAEL 59
 Db 25 VSPRLSGATSSSYGVFISNLR---KALPNKLYDIPLLRSSLPQSGQRYALIHITNYADE 81
 QY 60 SVTLALDVTNAYVGVYRAGNSAYFFHPDQCEDA-EALTHLFTVQVQRYTFAFGNVDRL 118
 Db 82 TTSVALDVTNAYVGVYRAGNSAYFFHPDQCEDA-EALTHLFTVQVQRYTFAFGNVDRL 138
 QY 119 QLAGNLRNIELNGPLLEAISALVYVSTGCTQPLTLARSFFIICQIMISPAARFQVIEGE 178
 Db 139 TAAGKIRENIPGLPALDSAITLIFYNAN-----SAASALMWLIQSTSEAAKYKLEQQ 193
 QY 179 MRTRI 183
 Db 194 IGRV 198

RESULT 4
 ABRC ABRPR STANDARD; PRT; 562 AA.
 AC P28590;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Abrin-c precursor [contains: Abrin-c A chain (rRNA N-glycosidase) (EC 3.2.2.22); Abrin-c B chain].
 OS Abrus precatorius (Indian licorice) (Crab's eye).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Abreace; Abrus.
 OX NCBI_TaxID=3816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf;
 RX MEDLINE=91266957; PubMed=2050149;
 RA Wood K.A., Lord J.M., Wawrzynczak E.J., Piatek M.;
 RT "Preproabrin: genomic cloning, characterisation and the expression of the A-chain in *Escherichia coli*.";
 RL Eur. J. Biochem. 198:723-732(1991).
 CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT PRECEDES ENDOCYTOSIS.
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC EMBL; X55667; CAA39202.1; --
 DR PIR; S16022; S16022.
 DR HSSP; P11140; IABR.
 DR InterPro; IPR000772; Ricin_B_lectin.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00652; Ricin_B_lectin; 6.
 DR Pfam; PF00361; RIP; 1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR SMART; SM00458; RICIN; 2.
 DR PROSITE; PS02311; RICIN B LECTIN; 2.
 DR PROSITE; PS0275; SHIGA-RICIN; 1.
 KW Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;

KW Glycoprotein; Lectin; Signal; Pyroglutamate carboxylic acid.
 FT SIGNAL 1 34 BY SIMILARITY.
 FT CHAIN 35 285 ABRIN C A CHAIN (BY SIMILARITY).
 FT PEPTIDE 286 295 LINKER PEPTIDE (BY SIMILARITY).
 FT CHAIN 296 562 ABRIN C B CHAIN (BY SIMILARITY).
 FT DOMAIN 307 434 RICIN B-TYPE LECTIN 1.
 FT DOMAIN 437 561 RICIN B-TYPE LECTIN 2.
 FT REPEAT 317 359 1-ALPHA.
 FT REPEAT 360 400 1-BETA.
 FT REPEAT 403 435 1-GAMMA.
 FT REPEAT 448 483 2-ALPHA.
 FT REPEAT 487 526 2-BETA.
 FT REPEAT 529 562 2-GAMMA.
 FT ACT_SITE 198 198 BY SIMILARITY.
 FT DISULFID 281 303 INTERCHAIN (BY SIMILARITY).
 FT DISULFID 320 339 BY SIMILARITY.
 FT DISULFID 363 380 BY SIMILARITY.
 FT DISULFID 451 464 BY SIMILARITY.
 FT DISULFID 490 507 BY SIMILARITY.
 FT MOD_RES 35 35 PYROGLUTAMATE CARBOXYLIC ACID (BY SIMILARITY).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 395 395 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 435 435 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 562 AA; 62817 MW; 1FD0ABC7D7BA6278 CRC64;
 Query Match 33.7%; Score 327.5; DB 1; Length 562;
 Best Local Similarity 43.9%; Pred. No. 1.1e-32;
 Matches 82; Conservative 22; Mismatches 68; Indels 15; Gaps 4;
 QY 6 QYPLINFTTAGATVQSYTNFIRAVRGLT-----VLENRVGLPINQRFILVELSNAH 57
 Db 35 QDQVKEITTEGATSSQYKQFIEALRQELTGLIHIDIPVLPDPTTVEERNYITVELSEN 94
 QY 58 ELSVTLALDVTNAYVGVYRAGNSAYFFHPDQCEDAETHLFTDQVQRYTFAFGNVDRL 117
 Db 95 RESIEVGIDVTNAYVAYRAGNSQSYFL---RDAPASASTYLPFGTQ-RYSLRFDGSGYDL 150
 QY 118 EQLAGNLRNIELNGPLLEAISALVYVSTGCTQPLTLARSFFIICQIMISPAARFQVIEG 177
 Db 151 ERWAHQTRFEEISLGLQALTHAIS---FLRGSANDEKARTLIVIIQWASEAARYIISN 207
 QY 178 EMRTIR 184
 Db 208 RVGSIR 214
 RESULT 5
 RIPS TRIKI STANDARD; PRT; 289 AA.
 AC P24478;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ribosome-inactivating protein karasurin precursor (rRNA N-glycosidase) (EC 3.2.2.22).
 OS Trichosanthes kirilowii (Mongolian snake-gourd).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
 OX NCBI_TaxID=3677;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Root tuber;
 RX MEDLINE=97356562; PubMed=9212998;
 RA Mizukami H., Iida K., Kondo T., Ogihara Y.;
 RT "Cloning and bacterial expression of a gene encoding ribosome-inactivating proteins, karasurin-A and karasurin-C, from *Trichosanthes kirilowii* var. japonica";
 RL Biol. Pharm. Bull. 20:711-713 (1997).
 RN [2]
 RP SEQUENCE OF 24-270.
 RX MEDLINE=92005921; PubMed=1914000;

MEDLINE=931132798; PubMed=8421313;
Hung C.-H., Lee M.-C., Lee T.C., Lin J.-Y.;
"Primary structure of three distinct isoforms determined by cDNA
sequencing. Conservation and significance."
J. Mol. Biol. 229:263-267(1993).
[2]
SEQUENCE OF 260-527.
TISSUE=Seed.
MEDLINE=931169023; PubMed=7763422;
Kimura M., Sumazawa T., Funatsu G.;
"The complete amino acid sequences of the B-chains of abrin-a and
abrin-b, toxic proteins from the seeds of Abrus precatorius.";
Biosci. Biotechnol. Biochem. 57:166-169(1993).
!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
ABRIN-A IS MORE TOXIC THAN RICIN.
!- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
PRECEDES ENDOCYTOSIS.
!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
specific adenosine on the 28S rRNA.
!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
!- SIMILARITY: Contains 2 ricin B-type lectin domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; M98345; AAA32625.1; -.
PIR; S32430; S32430.
HSPG; P11140; IABR.
InterPro; IPRO00772; Ricin_B_lectin.
InterPro; IPRO01574; RIP.
Pfam; PF00652; Ricin_B_lectin; 6.
Pfam; PF00161; RIP; 1.
PRINTS; PR00396; SHIGARICIN.
SMART; SM00458; RICIN; 2.
PROSITE; PS50231; RICIN_B_LLECTIN; 2.
PROSITE; PS00275; SHIGA_RICIN; 1.
Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
Glycoprotein; Lectin; Pyrrolidone carboxylic acid.
CHAIN 1 250 ABRI-B A CHAIN.
PEPTIDE 251 260 LINKER PEPTIDE.
CHAIN 261 527 ABRI-B B CHAIN.
DOMAIN 272 399 RICIN B-TYPE LECTIN 1.
DOMAIN 402 526 RICIN B-TYPE LECTIN 2.
REPEAT 282 324 1-ALPHA.
REPEAT 325 365 1-BETA.
REPEAT 368 400 1-GAMMA.
REPEAT 413 448 2-ALPHA.
REPEAT 452 491 2-BETA.
REPEAT 494 527 2-GAMMA.
ACT_SITE 163 163 BY SIMILARITY.
DISULFD 246 268 INTERCHAIN (BY SIMILARITY').
DISULFD 285 304 BY SIMILARITY.
DISULFD 328 345 BY SIMILARITY.
DISULFD 416 429 BY SIMILARITY.
DISULFD 455 472 BY SIMILARITY.
MOD_RES 1 1 PYROLIDONE CARBOXYLIC ACID (BY
SIMILARITY').

CARBOHYD 110 110 N-LINKED (GLCNAC...) (POTENTIAL).
FT FT CARBOHYD 360 360 N-LINKED (GLCNAC...) (POTENTIAL).
FT FT CARBOHYD 400 400 N-LINKED (GLCNAC...) (POTENTIAL).
FT FT CONFLICT 282 282 N > D (TN REF. 2)
FT FT

FT CONFLICT 291 291 D -> N (IN REF. 2).
 FT CONFLICT 350 351 AE -> PQ (IN REF. 2).
 FT CONFLICT 378 378 L -> M (IN REF. 2).
 FT CONFLICT 426 426 L -> M (IN REF. 2).
 FT CONFLICT 428 428 Y -> D (IN REF. 2).
 FT CONFLICT 431 431 N -> S (IN REF. 2).
 FT CONFLICT 431 431 N -> S (IN REF. 2).
 FT CONFLICT 484 484 R -> K (IN REF. 2).
 FT CONFLICT 484 484 R -> K (IN REF. 2).
 FT CONFLICT 491 491 N -> S (IN REF. 2).
 FT CONFLICT 493 493 H -> Y (IN REF. 2).
 FT CONFLICT 502 502 R -> G (IN REF. 2).
 FT CONFLICT 509 509 E -> Q (IN REF. 2).
 FT CONFLICT 513 513 H -> W (IN REF. 2).
 FT CONFLICT 516 516 H -> T (IN REF. 2).
 FT CONFLICT 527 527 AA; 59114 MW; 3253AE490C9494A CRC64;
 SQ SEQUENCE 527 AA; 59114 MW; 3253AE490C9494A CRC64;
 Query Match 33.3%; Score 323; DB 1; Length 527;
 Best Local Similarity 43.9%; Pred. No. 2.6e-22;
 Matches 83; Conservative 21; Mismatches 69; Indels 16; Gaps 4;
 QY 6 QYPLINFTAGATVQSYNTIRAVRGRLT-----VLNRVGLPINQRFILVELSNHA 57
 Db 1 QDQVIFKFTGATGSSQYKFIKALQRLTGGLIHGIPVLPDPTTLQBRNRYISVELSNSD 60
 QY 58 ELSVLTALDVNNAVGVYRAGNSAYFPHDPNQEDAEATHLFTDVQNRVTFAGGNYDRL 117
 Db 61 TESIEAGIDVSNAYVAYRAGNSYFL---RDAPTASRYLFTGTQ-QYSLRFGNSYIDL 116
 QY 118 EQLAGNIRENIEGNGPLEBAISALYYSTGGTQPLTARSFICIMISEAARFOYIEG 177
 Db 117 ERLAROTROQIPLGLQALRAHISFL-----QSGTDDQEIARTLIVIQWASEAARYRFSY 172
 QY 178 EMRTIRYN 186
 Db 173 RVGSIKTN 181

RESULT 7

RIP2_BRYDI STANDARD; PRT; 282 AA.
 AC P98184; Q9S8J0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ribosome-inactivating protein bryodin II precursor (rRNA N-glycosidase) (EC 3.2.2.22) (SD2).
 OS Bryonia dioica (Red Bryonia).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.
 OX NCBI_TaxID=3652;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Siegal C.B., Gawlak S.L., Marquardt H.;
 RT "Bryodin 2 a ribosome-inactivating protein isolated from the plant Bryonia dioica.";
 RL Bryonia dioica.";
 RN Patient number US5597569, 28-JAN-1997.
 RN [2]
 RP SEQUENCE OF 22-42.
 RC TISSUE=Root;
 RX MEDLINE=95151812; PubMed=7849072;
 RA Siegal C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
 RA Marquardt H.;
 RT "Characterization of ribosome-inactivating proteins isolated from Bryonia dioica and their utility as carcinoma-reactive immunocjugates.";
 RL Bioconj. Chem. 5:423-429(1994).
 CC -1- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS PROTEIN SYNTHESIS IN ANIMAL CELLS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one specific adenosine on the 28S rRNA.
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY. TYPE 1 RIP SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----NOT ANNOTATED_CDS.
 DR EMBL; I34238; -; NOT ANNOTATED_CDS.
 DR HSSP; P09989; 1MRJ.
 DR InterPro; IPR001574; RIP.
 DR Pfam; PF00161; RIP.1.
 DR PRINTS; PR00396; SHIGARICIN.
 DR PROSITE; PS00295; SHIGA_RICIN; 1.
 KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin;
 KW Multigene family; Glycoprotein; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 282 RIBOSOME-INACTIVATING PROTEIN BRYODIN II.
 FT ACT SITE 183 183 BY SIMILARITY.
 FT CARBOHYD 25 25 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 282 AA; 30754 MW; C52BE2F6A873769C CRC64;
 Query Match 33.0%; Score 320.5; DB 1; Length 282;
 Best Local Similarity 46.0%; Pred. No. 2.1e-22;
 Matches 81; Conservative 24; Mismatches 54; Indels 17; Gaps 7;
 QY 10 INETTATGATVQSYNTIRAVRGRLTV-LPNRVGLPINQ-----RFLVELSNHAELSV 61
 Db 24 INFSLGATGATYKTFIRNLRTKLTGTPRVVDIPVLRNAAAGLARFQLVLTNYNGESV 83
 QY 62 TLALDVNNAVGVYRAGNSAYFPHDPNQEDAEATHLFTDVQNRVTFAGGNYDRLQOLA 121
 Db 84 TVALDVNNAVGVYRAGNSAYFPHDPNQEDAEATHLFTDVQNRVTFAGGNYDRLQOLA 139
 QY 122 GNL-RENIENGPLEBAISALYYSTGGTQPLTARSFICIMISEAARFOYIE 176
 Db 140 GRISRENIENGPLEBAISALYYSTGGTQPLTARSFICIMISEAARFOYIE 191

RESULT 8

NIGB_SAMNI STANDARD; PRT; 563 AA.
 AC P33183; P33184; P93542;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Nigrin b precursor (Agglutinin V) (SNAV) [Contains: Nigrin b A chain (rRNA N-glycosidase) (EC 3.2.2.22); Nigrin b B chain].
 OS Sambucus nigra (European elder).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Campanulids; Dipsacales; Adoxaceae; Sambucus.
 OX NCBI_TaxID=4202;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Barb;
 RX MEDLINE=96215449; PubMed=8647092;
 RA Van Damme E.J., Barre A., Rouge P., Van Leuven F., Peumans W.J.;
 RT "Characterization and molecular cloning of Sambucus nigra agglutinin V (nigrin b), a GalNAC-specific type-2 ribosome-inactivating protein from the bark of elderberry (Sambucus nigra).";
 RL Eur. J. Biochem. 237:505-513(1996).
 RN [2]
 RP SEQUENCE OF 26-49 AND 298-321.
 RC TISSUE=Barb;
 RX MEDLINE=94003077; PubMed=8400135;
 RA Girbes T., Citores L., Ferreras J.M., Rojo M.A., Iglesias R., Munoz R., Arias F.J., Calonge M., Garcia J.R., Mendez E.;
 RT "Isolation and partial characterization of nigrin b, a non-toxic novel type 2 ribosome-inactivating protein from the bark of Sambucus nigra L.";
 RL Plant Mol. Biol. 22:1181-1186(1993).
 CC -1- FUNCTION: NON-TOXIC TYPE 2 RIP WHICH STRONGLY INHIBITS MAMMALIAN

CC PROTEIN SYNTHESIS BUT DOES NOT AFFECT PLANT NOR BACTERIAL PROTEIN
 CC SYNTHESIS. THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
 CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
 CC SUBUNITS BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA. THE
 CC B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT FACILITATES THE
 CC BINDING OF NIGRIN B TO THE CELL MEMBRANE THAT PRECEDES
 CC ENOCYTOSIS.
 CC
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE RIBOSOME-
 CC INACTIVATING PROTEIN FAMILY. TYPE 2 RIP SUBFAMILY.
 CC -!- SIMILARITY: Contains 2 ricin B-type lectin domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U41299; AAB39475.1; -;
 CC PIR; S37382; S37382.
 CC InterPro; IPR000772; Ricin_B_lectin.
 CC InterPro; IPR001574; RIP.
 CC Pfam; PF00652; Ricin_B_lectin; 6.
 CC Pfam; PF00161; RIP; 1.
 CC PRINTS; PR00396; SHIGARICIN.
 CC SMART; SM00458; RICIN; 2.
 CC PROSITE; PS00275; SHIGA_RICIN; 1.
 CC PROSITE; PS50231; RICIN_B_LECTIN; 2.
 CC Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 CC Glycoprotein; Lectin; Signal.
 CC SIGNAL 1 25
 CC CHAIN 26 297 NIGRIN B A CHAIN.
 CC CHAIN 298 563 NIGRIN B B CHAIN.
 CC DOMAIN 305 431 RICIN B-TYPE LECTIN 1.
 CC DOMAIN 434 559 RICIN B-TYPE LECTIN 2.
 CC REPEAT 316 356 1-ALPHA.
 CC REPEAT 357 397 1-BETA.
 CC REPEAT 400 432 1-GAMMA.
 CC REPEAT 445 482 2-BETA.
 CC REPEAT 486 524 2-ALPHA.
 CC REPEAT 527 554 2-GAMMA.
 CC ACT_SITE 188 188 INTERCHAIN (BY SIMILARITY).
 CC DISULFID 274 302 BY SIMILARITY.
 CC DISULFID 319 338 BY SIMILARITY.
 CC DISULFID 360 377 BY SIMILARITY.
 CC DISULFID 448 463 BY SIMILARITY.
 CC DISULFID 489 506 BY SIMILARITY.
 CC CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CONFLICT 39 39 K -> V (IN REF. 2).
 CC SEQUENCE 563 AA; 62300 MW; F250CB24621BF14 CRG64;
 CC
 CC Query Match 31.7%; Score 307.5; DB 1; Length 563;
 CC Best Local Similarity 38.4%; Pred. No. 7.6e-21;
 CC Matches 73; Conservative 36; Mismatches 56; Indels 25; Gaps 5;
 CC
 CC QY 7 YPIINFTAGATVQSYTNFIRAVR-----GRITVLPKRVGLPINOQFLVLSNH 56
 CC DB 28 YPVSVNLDGKATSYDIFLSNRKIVATGTVEVNGLPVLRRESEYQVKSRFVLVLTNY 87
 CC QY 57 AELSVTLALDVNTAYVGVYVGRAGNVAFFHPDNOEDAEAI--THLFDVQNRVTFAGGNY 114
 CC DB 88 NGNTVTLAVDVNTLYVAVSGNANSTFF-----KDTEVQKSNLFGVTGKQN-TLSTGNY 141
 CC QY 115 DRLEQLAGNLRNTELGNGLEPAISALYYSTGGTQLFPLARSFFICIMISEAARFOY 174

DB 142 DNLETAANTRRSEIELGSPLDGAITSLYHGD-----SVARSLLVVIQMVSEARFRY 194
 QY 175 IEQEMTRIR 184
 DB 195 IEQEVRSLSLQ 204
 CC
 CC RESULT 9
 CC RIPI_BRYDI
 CC ID RIPI_BRYDI STANDARD; PRT; 290 AA.
 CC AC P33185; O9S819;
 CC DT 01-OCT-1993 (Rel. 27, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Ribosome-inactivating protein bryodin I precursor (rRNA N-glycosidase)
 CC (EC 3.2.2.22) (BD1).
 CC OS Bryonia dioica (Red bryony).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
 CC OC eurosids I; Cucurbitales; Cucurbitaceae; Bryonia.
 CC OX NCBI_TaxID=3652;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 CC TISSUE=Leaf;
 CC RX MEDLINE=97228081; PubMed=9115985;
 CC RA Gawlak S.L., Neubauer M., Klei H.E., Chang C.Y.Y., Einspahr H.M.,
 CC Siegall C.B.;
 CC RT "Molecular, biological, and preliminary structural analysis of
 CC recombinant bryodin I, a ribosome-inactivating protein from the plant
 CC Bryonia dioica.";
 CC RL Biochemistry 36:3095-3103 (1997).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RA Siegall C.B.;
 CC RT "Cloning and expression of a gene encoding bryodin I from Bryonia
 CC dioica.";
 CC RL Patent number US5541110, 30-JUL-1996.
 CC RN [3]
 CC RP SEQUENCE OF 24-66.
 CC RC TISSUE=Seed;
 CC RX MEDLINE=89326691; PubMed=2753596;
 CC RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
 CC Lappi D.;
 CC RT "N-terminal sequence of some ribosome-inactivating proteins.";
 CC RL Int. J. Pept. Protein Res. 33:263-267 (1989).
 CC RN [4]
 CC RP SEQUENCE OF 24-43.
 CC RC TISSUE=Root;
 CC RX MEDLINE=95151812; PubMed=7849072;
 CC RA Siegall C.B., Gawlak S.L., Chace D., Wolff E.A., Mixan B.,
 CC Marquardt H.;
 CC RT "Characterization of ribosome-inactivating proteins isolated from
 CC Bryonia dioica and their utility as carcinoma-reactive
 CC immunoconjugates.";
 CC RL Bioconj. Chem. 5:423-429 (1994).
 CC CC -!- FUNCTION: RIBOSOME-INACTIVATING PROTEIN OF TYPE 1, INHIBITS
 CC PROTEIN SYNTHESIS IN ANIMAL CELLS.
 CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC -!- PTM: APPEARS TO UNDERGO PROTEOLYTIC CLEAVAGE IN THE C-TERMINAL TO
 CC PRODUCE A SHORTER PROTEIN.
 CC -!- BIOTECHNOLOGY: Especially useful as immunotoxin for
 CC pharmacological applications as it has low toxicity in rats and
 CC mice but is potent once inside target cells.
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 CC TYPE 1 RIP SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U41299; AAB39475.1; -;
 CC PIR; S37382; S37382.
 CC InterPro; IPR000772; Ricin_B_lectin.
 CC InterPro; IPR001574; RIP.
 CC Pfam; PF00652; Ricin_B_lectin; 6.
 CC Pfam; PF00161; RIP; 1.
 CC PRINTS; PR00396; SHIGARICIN.
 CC SMART; SM00458; RICIN; 2.
 CC PROSITE; PS00275; SHIGA_RICIN; 1.
 CC PROSITE; PS50231; RICIN_B_LECTIN; 2.
 CC Plant defense; Hydrolase; Protein synthesis inhibitor; Toxin; Repeat;
 CC Glycoprotein; Lectin; Signal.
 CC SIGNAL 1 25
 CC CHAIN 26 297 NIGRIN B A CHAIN.
 CC CHAIN 298 563 NIGRIN B B CHAIN.
 CC DOMAIN 305 431 RICIN B-TYPE LECTIN 1.
 CC DOMAIN 434 559 RICIN B-TYPE LECTIN 2.
 CC REPEAT 316 356 1-ALPHA.
 CC REPEAT 357 397 1-BETA.
 CC REPEAT 400 432 1-GAMMA.
 CC REPEAT 445 482 2-BETA.
 CC REPEAT 486 524 2-ALPHA.
 CC REPEAT 527 554 2-GAMMA.
 CC ACT_SITE 188 188 INTERCHAIN (BY SIMILARITY).
 CC DISULFID 274 302 BY SIMILARITY.
 CC DISULFID 319 338 BY SIMILARITY.
 CC DISULFID 360 377 BY SIMILARITY.
 CC DISULFID 448 463 BY SIMILARITY.
 CC DISULFID 489 506 BY SIMILARITY.
 CC CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 537 537 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CONFLICT 39 39 K -> V (IN REF. 2).
 CC SEQUENCE 563 AA; 62300 MW; F250CB24621BF14 CRG64;
 CC
 CC Query Match 31.7%; Score 307.5; DB 1; Length 563;
 CC Best Local Similarity 38.4%; Pred. No. 7.6e-21;
 CC Matches 73; Conservative 36; Mismatches 56; Indels 25; Gaps 5;
 CC
 CC QY 7 YPIINFTAGATVQSYTNFIRAVR-----GRITVLPKRVGLPINOQFLVLSNH 56
 CC DB 28 YPVSVNLDGKATSYDIFLSNRKIVATGTVEVNGLPVLRRESEYQVKSRFVLVLTNY 87
 CC QY 57 AELSVTLALDVNTAYVGVYVGRAGNVAFFHPDNOEDAEAI--THLFDVQNRVTFAGGNY 114
 CC DB 88 NGNTVTLAVDVNTLYVAVSGNANSTFF-----KDTEVQKSNLFGVTGKQN-TLSTGNY 141
 CC QY 115 DRLEQLAGNLRNTELGNGLEPAISALYYSTGGTQLFPLARSFFICIMISEAARFOY 174

Query Match 31.3%; Score 303.5; DB 1; Length 528;
Best Local Similarity 41.1%; Pred. No. 1.6e-20;
Matches 76; Conservative 26; Mismatches 64; Indels 19


```
R1PA_LURCY
ID R1PA_LURCY STANDARD; PRT; 277 AA.
AC Q00465;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribosome-inactivating protein luffin-alpha precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22)
OS Luffa cylindrica (Smooth loofah) (Sponge gourd).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3670;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=92288316; PubMed=1600156;
RA Kataoka J., Habuka N., Miyano M., Masuta C., Koizumi A.;
RT "Nucleotide sequence of cDNA encoding alpha-luffin, a ribosome-
RT inactivating protein from luffa cylindrica.";
RL Plant Mol. Biol. 18:1199-1202(1992).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X62371; CAA44229.1; -
DR PIR; S22494; S22494.
DR HSP; P16094; 1AHC.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PRO0396; SHIGARICIN.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
KW Plant defense; Antiviral; Protein synthesis inhibitor; Hydrolase;
KW Toxin; Signal.
FT SIGNAL 1 19
FT CHAIN 20 277 RIBOSOME-INACTIVATING PROTEIN LUFFIN-
FT ALPHA.
FT ACT_SITE 179 179 BY SIMILARITY.
FT SEQUENCE 277 AA; 30212 MW; EA17FC27998C25AC CRC64;
SQ
Query Match 28.3%; Score 274.5; DB 1; Length 277;
Best Local Similarity 33.3%; Pred. No. 3.6e-18;
Matches 63; Conservative 42; Mismatches 61; Indels 23; Gaps 5;
QY 10 INFTTAGATVQSYTNFIRAVRGRGLTVLPNRVGLPIN-----QRFILVELSNHAE 58
DB 22 VRFSLGSSSTVSXKFIGDLR---KALPN-GTVYNTILLSSAGASRYTLWLSNYDG 77
QY 59 LSVTLALDVTNAYVGYRAGNSAYFFHPDQDAEAITLFTDVQVRYTFAGGNYDRLE 118
DB 78 KAITVAVDVTNVMGYLVNYSYFF---NESDAKLASQYVFGKSTIVLPYSGNYEKLU 134
QY 119 QLAGNLRNRELNGPLERAIASLYYYTGGLTGLPLARSFFICIMISEAARFOYIEGE 178
DB 135 TAAGKIREKIPGLFPALDSAITLTFHYDS-----TAAAAAFVLIOTTAEGRFKYIEQ 189
QY 179 MTRIRYNR 187
DB 190 IIRISKNQ 198
RESULT 14
ID R1PI_MOMCH STANDARD; PRT; 286 AA.
```

```
AC P16094; P24697;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ribosome-inactivating protein momordin I precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22) (Alpha-momorcharin) (Alpha-MMC).
OS Momordica charantia (Bitter melon) (Balsam pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3673;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=91159486; PubMed=2001404;
RA Ho W.K.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
RT "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating
RT protein.";
RL Biochim. Biophys. Acta 1088:311-314(1991).
RN [2]
RP SEQUENCE OF 24-38.
RC TISSUE=Seed;
RX MEDLINE=89326691; PubMed=2753596;
RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
RA Lappi D.;
RT "N-terminal sequence of some ribosome-inactivating proteins.";
RL Int. J. Pept. Protein Res. 33:263-267(1989).
RN [3]
RP SEQUENCE OF 24-70.
RC TISSUE=Seed;
RX MEDLINE=89005108; PubMed=3262509;
RA Casellas P., Dussosoy D., Palasca A.I., Barbieri L.,
RA Guilleminot J.C., Ferrara P., Bolognesi A., Cennini P., Stirpe F.;
RT "Trichokirin, a ribosome-inactivating protein from the seeds of
RT Trichosanthes kirilowii Maximowicz. Purification, partial
RT characterization and use for preparation of immunotoxins.";
RL Eur. J. Biochem. 176:581-588(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX MEDLINE=94356447; PubMed=8075985;
RA Ren J., Wang Y., Dong Y., Stuart D.I.;
RT "The N-glycosidase mechanism of ribosome-inactivating proteins
RT implied by crystal structures of alpha-momorcharin.";
RL Structure 2:7-16(1994).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
RX MEDLINE=94192822; PubMed=8143869;
RA Husain J., Tickle I.J., Wood S.P.;
RT "Crystal structure of momordin, a type I ribosome inactivating
RT protein from the seeds of Momordica charantia.";
RL FEBS Lett. 342:154-158(1994).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=95344383; PubMed=7619070;
RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
RT "Studies on crystal structures, active-centre geometry and
RT depurinating mechanism of two ribosome-inactivating proteins.";
RL Biochem. J. 309:285-298(1995).
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X57682; CAA40869.1; -
DR PIR; S14273; RLPUGG.
```

```

QY      122 GNLRNIELNGPLEPAISALYYVSTGGTQLPTLARSFICIQMISEARFYIEGEMRT 181
       ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB     142 GKXPEKIPIGLPALDSAISTLLHYDS-----TAAGALLVLQTAAEAARKFYEQIOQE 196

QY      182 R 182
DB     197 R 197

RESULT 15
RIP2 MOWBA          STANDARD;           PRT;    286 AA.
AC   E29339;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DE   15-SEP-2003 (Rel. 42, Last annotation update)
DR   Ribosome-inactivating protein momordin II precursor (rRNA
DE N-glycosidase) (EC 3.2.2.22).
OS Momordica balsamina (Bitter melon) (Balsam pear).
OC Caryophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Momordica.
OX NCBI_TaxID=3672;
RN [1]
SEQUENCE FROM N.A.
RP RP
RC TISSUE=Seed;
RX MEDLINE=93027170; PubMed=1408771;
RA Ortigao M., Better M.;
RT "Momordin II, a ribosome inactivating protein from Momordica
RT balsamina, is homologous to other plant proteins.";
RL Nucleic Acids Res. 20:4662-4662(1992).
CC -I- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one
CC specific adenosine on the 28S rRNA.
CC -I- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
CC TYPE 1 RIP SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as it's content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Z12175; CAAT8166.1; -.
CC FJX; S25560; S25560.
CC DB; ICES; 07-JUN-99.
CC InterPro; IPRO01574; RIP.
CC Pfam; PF00161; Rip; 1.
CC PRINTS; PR00396; SHIGARICIN.
CC PROSITE; PS00275; SHIGA_RICIN; 1.
CC Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin; Signal;
CC 3D-structure.
FT SIGNAL                     1 23
FT CHAIN                      24 286        RIBOSOME-INACTIVATING PROTEIN MOMORDIN
FT FT                          BY SIMILARITY.
FT ACT SITE                   181         181
SQ SEQUENCE                    286 AA; 32031 MW; 3B89FFFAE6B25986 CRC64;
Query Match               28.0%; Score 272; DB 1; Length 286;
Best Local Similarity    34.2%; Pred. No. 6.3e-18;
Matches                  63; Conservative            39; Mismatches             66; Indels              16; Gaps                4;

QY      10 INFTTAGATVCQSNTFNIRAVRGRLTLVLPNRVGUPI-----NORFLVELSNHAELSVT 62
       ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB     25 VNFDSLSTATAKTKIKIEDPRALTPLPSHKYVDIIPLYSTSDSRFRILLDLTSVAYETIS 84

QY      63 LALDVTNAYVVGYRAGNSAIFYFHFDNQEDAETHLFVTQVRNYTFYAFGNYSRLEOLAG 122
       ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB     85 VALDIVNVVVAVTRDVDVSFFF--KESPPEAYNILFKGR-KITLYPYTGNYENLQTPAH 140

QY      123 NLRENIENGCPLEBASIALYYYSTGTQLPTLARSPFIQMISEARFYIEGEMRTR 182
       ||||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
```

Db 141 KIRENID/LPALSSAITTLFYNA-----QSAPSALLVLIOTTAARFKYIERHVAKY 195

Qy 183 IRYN 186

Db :
196 VATN 199

Search completed: February 10, 2004, 16:23:24
Job time : 7.30467 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:17:00 ; Search time 25.2187 Seconds
(without alignments)
1933.961 Million cell updates/sec

Title: US-10-083-336A-6

Perfect score: 971

Sequence: 1 MTFPKQYPIINTTAGATVQ.....ARFQYIEGEMTRIRYNRRS 189

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

1: sp arChaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_arChaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	951	97.9	541	10 Q41174	Q41174 ricinus com
2	374.5	38.6	580	10 Q94BW3	Q94BW3 cinnamomum
3	372.5	38.4	580	10 Q94BW4	Q94BW4 cinnamomum
4	370.5	38.2	581	10 Q94BW5	Q94BW5 cinnamomum
5	368.5	38.0	549	10 Q95V22	Q95V22 cinnamomum
6	334	34.4	289	10 Q94KE4	Q94KE4 trichosanath
7	332	34.2	289	10 Q41216	Q41216 trichosanath
8	329.5	33.9	563	10 Q04367	Q04367 sambucus ni
9	326	33.6	247	10 Q9LRE3	Q9LRE3 trichosanath
10	323.5	33.3	564	10 Q9AVR2	Q9AVR2 sambucus eb
11	318.5	32.8	528	10 Q06076	Q06076 abrus preca
12	317	32.6	270	10 Q8LPV7	Q8LPV7 trichosanath
13	310.5	32.0	252	10 Q38760	Q38760 abrus preca
14	307.5	31.7	563	10 Q945S2	Q945S2 sambucus ni
15	307.5	31.7	563	10 Q8GT32	Q8GT32 sambucus ni
16	305	31.4	270	10 Q41611	Q41611 trichosanath

17	303.5	31.3	252	10	Q38761	Q38761 abrus preca
18	302	31.1	565	10	Q04071	Q04071 sambucus ni
19	300.5	30.9	278	10	Q00980	Q00980 luffa cylin
20	298.5	30.7	547	10	Q9M6E9	Q9M6E9 abrus preca
21	297.5	30.6	251	10	Q96236	Q96236 abrus preca
22	296.5	30.5	251	10	Q96237	Q96237 abrus preca
23	293.5	30.2	566	10	Q04072	Q04072 sambucus ni
24	290.5	29.9	251	10	Q96235	Q96235 abrus preca
25	273	28.1	264	10	Q9FSH2	Q9FSH2 momordica c
26	273	28.1	570	10	Q41358	Q41358 sambucus ni
27	271	27.9	592	10	Q8W2E7	Q8W2E7 iris hollan
28	270	27.8	249	10	Q8LK05	Q8LK05 viscum albu
29	270	27.8	286	10	Q9FUV7	Q9FUV7 momordica c
30	267	27.5	570	10	Q22415	Q22415 sambucus ni
31	266.5	27.4	604	10	Q9M654	Q9M654 polygonatum
32	265	27.3	286	10	Q41257	Q41257 momordica c
33	263	27.1	565	10	Q8W243	Q8W243 viscum albu
34	260	26.8	254	10	Q8LK06	Q8LK06 viscum albu
35	260	26.8	573	10	Q8W2E8	Q8W2E8 iris hollan
36	259	26.7	251	10	Q8LKQ4	Q8LKQ4 viscum albu
37	259	26.7	293	10	Q8S452	Q8S452 jatrophia cu
38	259	26.7	569	10	Q93543	Q93543 sambucus ni
39	253	26.1	531	10	Q8RXH6	Q8RXH6 viscum albu
40	251	25.8	249	10	Q8RXH7	Q8RXH7 viscum albu
41	247	25.4	258	10	Q9S9E4	Q9S9E4 gelonium mu
42	247	25.4	293	10	Q8VYU0	Q8VYU0 jatrophia cu
43	246.5	25.4	203	10	Q8RY69	Q8RY69 gynostemma
44	246.5	25.4	275	10	Q8H1Y4	Q8H1Y4 gynostemma
45	245.5	25.3	603	10	Q9M653	Q9M653 polygonatum

ALIGNMENTS

RESULT 1

Q41174
ID Q41174 PRELIMINARY; PRT; 541 AA.
AC Q41174;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Proricin A chain (EC 3.2.2.22) (rRNA N-glycosidase)
DE (Fragment)
OS Ricinus communis (Castor bean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Malpighiales; Euphorbiaceae; Ricinus.
OX NCBI_TaxID=3988;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92338377; PubMed=1633311;
RA Roberts L.M., Tregear J.W., Lord J.M.;
RT "Molecular cloning of ricin."
RL Targeted Diagn. Ther. 7:81-97(1992).
CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; S40366; AAB22582.1; -.
DR HSSP; P02879; 1BR6.
DR InterPro; IPR000172; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR InterPro; IPR001400; Somatotropin.
DR Pfam; PF00852; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS050231; RICIN_B_LECTIN; 2.
DR PROSITE; PS00275; SHIGA_RICIN; 1.
DR PROSITE; PS00338; SOMATOTROPIN_2; 1.
KW Hydrolase; Toxin.
FT NON TER 1
SQ SEQUENCE 541 AA; 60281 MW; 2B7B2CDF1F2E9D9 CRC64;

```
Query Match      97.9%; Score 951; DB 10; Length 541;
Best Local Similarity 94.9%; Pred. No. 2.8e-82;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY      2  IFKQYPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINOREFLV 51
DB      1  IFKQYPIINFNTAGATVQSYTNFIRAVRGRLTGGADVRHDIPVLPNRVGLPINOREFLV 60

QY      52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYTFAG 111
DB      61  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYTFAG 120

QY      112  GNYDELEQAGLNRENIELNGPLAEASALYYSTGGTQPLTARSFIICQMISEAR 171
DB      121  GNYDELEQAGLNRENIELNGPLAEASALYYSTGGTQPLTARSFIICQMISEAR 180

QY      172  FOYIEGEMTRIRYNRRS 189
DB      181  FOYIEGEMTRIRYNRRS 198

RESULT 2
Q94BW3
ID Q94BW3 PRELIMINARY; PRT; 580 AA.
AC Q94BW3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin III precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DE EMBL; AY039803; AAK82460.1; -
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32
FT CHAIN 33 580
FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN III.
SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7F9558 CRC64;

Query Match      38.6%; Score 374.5; DB 10; Length 580;
Best Local Similarity 48.1%; Pred. No. 3.7e-27;
Matches 90; Conservative 26; Mismatches 56; Indels 15; Gaps 5;

QY      7  YPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINOREFLVLSN-H 56
DB      33  YQTVFTTKNATKTSYQFIEALRAQLASGEHPGIPWNRDSTVPDSKRFLVLSNWA 92

QY      57  AELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYTFAGGNYDR 116
DB      93  ADSPVTLADVNTNAYVAVRTGSGFFLRDNPDP--PAIENLLPTK-RYTFPFSSGYTD 149

QY      117  LEQAGLNRENIELNGPLAEASALYYSTGGTQPLTARSFIICQMISEARFOVIE 176
DB      150  LERVAGERREILLGMDPLENAISALWISNL--NQORALARSLIVVQWVAEAVRFRFIE 207

Query Match      38.4%; Score 372.5; DB 10; Length 580;
Best Local Similarity 47.6%; Pred. No. 5.8e-27;
Matches 89; Conservative 27; Mismatches 56; Indels 15; Gaps 5;

QY      7  YPIINFNTAGATVQSYTNFIRAVRGRLTVPNRVGLPT-----NORFILVELSN-H 56
DB      33  YQTVFTTKNATKTSYQFIEALRAQLASGEHPGIPWNRDSTVPDSKRFLVLSNWA 92

QY      57  AELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYTFAGGNYDR 116
DB      93  ADSPVTLADVNTNAYVAVRTGSGFFLRDNPDP--PAIENLLPTK-RYTFPFSSGYTD 149

QY      117  LEQAGLNRENIELNGPLAEASALYYSTGGTQPLTARSFIICQMISEARFOVIE 176
DB      150  LERVAGERREILLGMDPLENAISALWISNL--WTSNLTQQORALARSLIVVQWVAEAVRFRFIE 207

RESULT 4
Q94BW5
ID Q94BW5 PRELIMINARY; PRT; 581 AA.
AC Q94BW5
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin I precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
```

```
Query Match      97.9%; Score 951; DB 10; Length 541;
Best Local Similarity 94.9%; Pred. No. 2.8e-82;
Matches 188; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

QY      2  IFKQYPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINOREFLV 51
DB      1  IFKQYPIINFNTAGATVQSYTNFIRAVRGRLTGGADVRHDIPVLPNRVGLPINOREFLV 60

QY      52  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYTFAG 111
DB      61  ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYTFAG 120

QY      112  GNYDELEQAGLNRENIELNGPLAEASALYYSTGGTQPLTARSFIICQMISEAR 171
DB      121  GNYDELEQAGLNRENIELNGPLAEASALYYSTGGTQPLTARSFIICQMISEAR 180

QY      172  FOYIEGEMTRIRYNRRS 189
DB      181  FOYIEGEMTRIRYNRRS 198

RESULT 2
Q94BW3
ID Q94BW3 PRELIMINARY; PRT; 580 AA.
AC Q94BW3
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type 2 ribosome-inactivating protein cinnamomin III precursor
DE (EC 3.2.2.22) (rRNA N-glycosidase).
OS Cinnamomum camphora (Camphor tree).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
OX NCBI_TaxID=13429;
RN [1]
RP SEQUENCE FROM N.A.
RA Yang Q., Gong Z.Z., Liu W.Y.;
RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
RT genes encoding cinnamomin proteins and study of their expression
RT patterns.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DE EMBL; AY039803; AAK82460.1; -
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS50231; RICIN_B_LECTIN; 2.
KW Hydrolase; Signal; Toxin.
FT SIGNAL 1 32
FT CHAIN 33 580
FT TYPE 2 RIBOSOME-INACTIVATING PROTEIN
FT CINNAMOMIN III.
SQ SEQUENCE 580 AA; 64421 MW; 940D10F01E7F9558 CRC64;

Query Match      38.6%; Score 374.5; DB 10; Length 580;
Best Local Similarity 48.1%; Pred. No. 3.7e-27;
Matches 90; Conservative 26; Mismatches 56; Indels 15; Gaps 5;

QY      7  YPIINFNTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINOREFLVLSN-H 56
DB      33  YQTVFTTKNATKTSYQFIEALRAQLASGEHPGIPWNRDSTVPDSKRFLVLSNWA 92

QY      57  AELSVTLALDVTNAYVVGVRAGNSAYFFHPDNQEDAEATHLFTDVQNRVYTFAGGNYDR 116
DB      93  ADSPVTLADVNTNAYVAVRTGSGFFLRDNPDP--PAIENLLPTK-RYTFPFSSGYTD 149

QY      117  LEQAGLNRENIELNGPLAEASALYYSTGGTQPLTARSFIICQMISEARFOVIE 176
DB      150  LERVAGERREILLGMDPLENAISALWISNL--NQORALARSLIVVQWVAEAVRFRFIE 207
```

OS Cinnamomum camphora (Camphor tree).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
 OX NCBI_TaxID=13429;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yang Q., Gong Z.Z., Liu W.Y.;
 RT "Molecular cloning of three type 2 RIP (ribosome-inactivating protein)
 RT genes encoding cinnamomin proteins and study of their expression
 RT patterns.";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 DR EMBL: AY039801; AAK82458.1; --
 DR InterPro: IPR000772; Ricin_B_lectin.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00652; Ricin_B_lectin; 5.
 DR Pfam: PF00161; RIP; 1.
 DR PRINTS: PR00396; SHIGARICIN.
 DR SMART: SM00458; RICIN; 2.
 DR PROSITE: PS0231; RICIN_B_LECTIN; 2.
 KW Hydrolase; Signal; Toxin.
 FT SIGNAL 1 32
 FT CHAIN 33 581
 FT CINNAMOMIN I.
 FT SEQUENCE 581 AA; 64215 MW; 6E8F5FB8A3D196 CRC64;
 Query Match 39.2%; Score 370.5; DB 10; Length 581;
 Best Local Similarity 48.1%; Pred. No. 9e-27;
 Matches 90; Conservative 25; Mismatches 57; Indels 15; Gaps 5;
 QY 7 YPIINFTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFILLVELSN-H 56
 Db 33 YQVTFTTKKATKTSYQFIEALRAQLASGEHPHGIPVMRERSTVPDSKRFILVELSNWA 92
 QY 57 AELSVTALDVTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFGNYDR 116
 Db 93 ADSPTVLAVDVTNAYVAYVYTGSGFLEDRNPD--PAIENLLPDYK-RYTFPFGSGYTD 149
 QY 117 LEQAGNLRNIELGNGPLEEAISALYYSTGTQTLPTLARSFIICIMISEAARFOYIE 176
 Db 150 LEGVAGERREILLGMDPLENAISALWISNL--NQORALARSILVVIQMAEAVRFRFIE 207
 QY 177 GEMRTI 183
 Db 208 YRVGSI 214
 RESULT 5
 ID Q9FV22 PRELIMINARY; PRT; 549 AA.
 AC Q9FV22
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Type II ribosome-inactivating protein cinnamomin (EC 3.2.2.22) (rRNA
 DE N-glycosidase) (Fragment).
 OS Cinnamomum camphora (Camphor tree).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Laurales; Lauraceae; Cinnamomum.
 OX NCBI_TaxID=13429;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Xie L., Liu W.-Y., Wang E.-D.;
 RT "Molecular cloning of cinnamomin A-, B-chain and the expression,
 RT purification, characterization and mutagenesis of the A-Chain.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 DR EMBL: AF259548; AAF68978.2; --
 DR HSP; P02879; 2AAI.

DR InterPro: IPR000772; Ricin_B_lectin.
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00652; Ricin_B_lectin; 5.
 DR Pfam: PF00161; RIP; 1.
 DR PRINTS: PR00396; SHIGARICIN.
 DR SMART: SM00458; RICIN; 2.
 DR PROSITE: PS0231; RICIN_B_LECTIN; 2.
 KW Hydrolase; Toxin.
 FT NON_TER 1
 FT SEQUENCE 549 AA; 60648 MW; 02607FE607CA44B0 CRC64;
 Query Match 38.0%; Score 368.5; DB 10; Length 549;
 Best Local Similarity 48.1%; Pred. No. 1.3e-26;
 Matches 90; Conservative 25; Mismatches 57; Indels 15; Gaps 5;
 QY 7 YPIINFTAGATVQSYTNFIRAVRGRLT-----VLPNRVGLPINORFILLVELSN-H 56
 Db 1 YQVTFTTKKATKTSYQFIEALRAQLASGEHPHGIPVMRERSTVPDSKRFILVELSNWA 60
 QY 57 AELSVTALDVTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFGNYDR 116
 Db 61 ADSPTVLAVDVTNAYVAYVYTGSGFLEDRNPD--PAIENLLPDYK-RYTFPFGSGYTD 117
 QY 117 LEQAGNLRNIELGNGPLEEAISALYYSTGTQTLPTLARSFIICIMISEAARFOYIE 176
 Db 118 LEGVAGERREILLGMDPLENAISALWISNL--NQORALARSILVVIQMAEAVRFRFIE 175
 QY 177 GEMRTI 183
 Db 176 YRVGSI 182
 RESULT 6
 ID Q94KE4 PRELIMINARY; PRT; 289 AA.
 AC Q94KE4
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Trichosanthin precursor (EC 3.2.2.22) (rRNA N-glycosidase).
 GN TCS.
 OS Trichosanthin kirilowii (Mongolian snake-gourd).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Cucurbitales; Cucurbitaceae; Trichosanthes.
 OX NCBI_TaxID=3677;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yuan H., Wang L., Wang Y., An C., Chen Z.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
 CC SPECIFIC ADENOSINE ON THE 28S RNA.
 CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
 DR EMBL: AF367252; AAK52960.1; --
 DR InterPro: IPR001574; RIP.
 DR Pfam: PF00161; RIP; 1.
 DR PRINTS: PR00396; SHIGARICIN.
 DR PROSITE: PS00275; SHIGA_RICIN; 1.
 KW Hydrolase; Signal; Toxin.
 FT SIGNAL 1 23
 FT CHAIN 24 270
 FT TRICHOSANTHIN.
 FT SEQUENCE 289 AA; 31706 MW; A6D5602549CA5657 CRC64;
 Query Match 34.4%; Score 334; DB 10; Length 289;
 Best Local Similarity 38.9%; Pred. No. 1.1e-23;
 Matches 72; Conservative 42; Mismatches 49; Indels 22; Gaps 5;
 QY 10 INFTTAGATVQSYTNFIRAVRGRLTVLPN-----RVGLPINORFILLVELSNHAE 59
 Db 25 VSFRLSGATSSSYGVFISNR---KALPNERKLYDIPLRSSLPFGSQRTALHILHTYADE 81
 QY 60 SVTLALDVTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFGNYDRLE 118
 DR EMBL: AF259548; AAF68978.2; --

SQ SEQUENCE 564 AA; 62694 MW; 8261681A6DB55CB8 CRC64;

Query Match 33.3%; Score 323.5; DB 10; Length 564;
Best Local Similarity 40.0%; Pred. No. 2.7e-22;
Matches 76; Conservative 38; Mismatches 51; Indels 25; Gaps 6

QY 7 YPIINFTTAGATVQSNTNFIRAVGR-----LTVLPRNRVGLPINORFILVELSNH 56
Db :
28 YPSVSFNLAGAKSTTYRDFLKNLRDRAVTGTVENGFLPVRRESEVQVKNRFLVRLTNY 87
QY AELSVTLALDVNAYVVGYRAGNSAYFFHPDNQEDAEAI--THLFTDVQNRYTFAPCGNY 114
Db :
88 NGDVTISADVNNILYLVA SANGNSYFF-----KDATELOKSNLFCT-QHTLSFTFGY 141
QY DRLEQLAGNLRENIELGNGLPEEAI SALYYSPGTGQTLP LARSFTIICIMISEARFOY 174
Db :
142 DNLETAAGTRRESIELGPNPLDGAITSLWY-DGG-----VARSLLVLIQMVEAFERY 194
QY I75 IECEMRTRIR 184
Db |||:||:|:
195 IEQEVRRLSQ 204

RESULT 11

Q06076 PRELIMINARY; PRT; 528 AA.

ID Q06076; AC Q06076; DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DS Abrin-d (EC 3.2.2.22) (rRNA N-glycosidase) (Fragment).
OS Abrus precatorius (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosoids I; Fabales; Fabaceae; Papilionoideae; Abreace; Abrus.
NCBI_TaxID=3816;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93132798; PubMed=8421313;
RT Hung C.-H., Lee M.-C., Lee T.-C., Lin J.-Y.;
RT "Primary structure of three distinct isoabrin determined by cDNA
RT sequencing: conservation and significance.";
RJ J. Mol. Biol. 229:263-267(1993).
CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC -!- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING PROTEIN FAMILY.
DR EMBL; X98346; AAA32626.1; -.
DR HSSP; P11140; LABR.
DR InterPro; IPR000772; Ricin_B_lectin.
DR InterPro; IPR001574; RIP.
DR Pfam; PF00652; Ricin_B_lectin; 6.
DR Pfam; PF00161; RIP; 1.
DR PRINTS; PR00396; SHIGARICIN.
DR SMART; SM00458; RICIN; 2.
DR PROSITE; PS02331; RICIN_B_LECTIN; 2.
DR PROSITE; PS0275; SHIGA_RICIN; 1.
DR Hydrolase; Toxin.
FT NON TER 1
FT NON TER 528
FT NON TER 528
SQ SEQUENCE 528 AA; 58870 MW; 62ED42FB8FFE60F8 CRC64;

Query Match 32.8%; Score 318.5; DB 10; Length 528;
Best Local Similarity 43.3%; Pred. No. 7.4e-22;
Matches 81; Conservative 22; Mismatches 69; Indels 15; Gaps 4;

QY 6 QYPIINFTTAGATVQSNTNFIRAVGRILT-----VLPRNVGLPINORFILVELSNHA 57
Db :
1 QDQVIKFTEGATSQSYKOFIEALRSRITGLIHDPVLPDPPTVEERNRYITVELSNS 60
QY 58 EELSVTLALDVNAYVVGYRAGNSAYFFHPDNQEDAEATHLFTDVQNRYTFAGGNYDL 117
Db :
61 RESIEVGIDVNAYVVA RAGSQSFV---RDAPASATYLEFPQTQ-RYSLEFDGSIGDL 116

OS *Abrus precatorius* (Indian licorice) (Crab's eye).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids 1; Fabales; Fabaceae; Papilionoideae; Abreae; Abrus.
NCBI_TaxID=3816;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=LEAF.
RC MEDLINE=91201329; PubMed=2016300;
RX Evensen G., Mathiesen A., Sundan A.;
RA "Direct molecular cloning and expression of two distinct abrin A-
RT chains".
RT J. Biol. Chem. 266:6848-6852(1991).
RL CC -!- FUNCTION: THE A CHAIN IS RESPONSIBLE FOR INHIBITING PROTEIN
CC SYNTHESIS THROUGH THE CATALYTIC INACTIVATION OF 60S RIBOSOMAL
CC SUBUNIT BY REMOVING ADENINE FROM POSITION 4,324 OF 28 S RNA.
CC ABRIN-A IS MORE TOXIC THAN RICIN.
CC CC -!- FUNCTION: THE B CHAIN IS A GALACTOSE-SPECIFIC LECTIN THAT
CC FACILITATES THE BINDING OF ABRIN TO THE CELL MEMBRANE THAT
CC PRECEDES ENDOCYTOSIS.
CC CC -!- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF THE N-GLYCOSIDIC BOND AT ONE
CC SPECIFIC ADENOSINE ON THE 28S RNA.
CC CC -!- SUBUNIT: DISULFIDE-LINKED DIMER OF A AND B CHAINS.
CC CC -!- DOMAIN: THE B CHAIN IS COMPOSED OF TWO DOMAINS, EACH DOMAIN
CC CONSISTS OF 3 HOMOLOGOUS SUBDOMAINS (ALPHA, BETA, GAMMA).
CC CC -!- SIMILARITY: TO OTHER BACTERIAL AND PLANTS RIBOSOME-INACTIVATING
CC PROTEINS. BELONGS TO TYPE 2 RIP.
CC EMBL; X54872; CAA38654.1; -;
CC HSSP; P11140; 1ABR.
CC InterPro; IPR001574; RIP.
CC Pfam; PF00161; RIP; 1.
CC PROSITE; PR00396; SHIGARICIN.
CC DR PROSITE; PS00275; SHIGA RICIN; 1.
CC DR Hydroxylase; Glycosidase; Toxin; Repeat; Glycoprotein; Lectin.
CC KW CHAIN 1 252 ABRIN B, A CHAIN (BY SIMILARITY).
CC FT NON_TER 252 252
CC FT SEQUENCE 252 AA; 28309 MW; BFFC846B9E92B5DE CRC64;
CC SQ
Query Match 32.0%; Score 310.5; DB 10; Length 252;
Best Local Similarity 43.2%; Pred. No. 1.6e-21;
Matches 79; Conservative 22; Mismatches 67; Indels 15; Gaps 4;
QY 10 INTTGAATVQSYTNFTRAVGRH-----TVLPNRVGLPINQRFILVELSNHAELSV 61
Db 6 IKFSTEGATQSQYKQFTEALRELRGLIHDPVLRDPTTVEERNRYITVELSNSRESI 65
QY 62 TLALDVNAVVGVRAGNSAYFFHPDQDAEALHTLFTDVQNRYYTFAFGNGYRLEOLA 121
Db 66 EVGIDVNAVVAYRAGSQSYEL--RDAPASATYLTGTQ-RYSLRFDGSGDLERWA 121
QY 122 GNRENIELNGNPIEAEISALYYVSTGCTQLPTLARSFITCIQMISARQYIEGENMT 181
Db 122 HQTREQISLGLOALTHAIS---FLRSGASNDDEKARTLIVIOAWSEAAARYRYSNRVGV 178
QY 182 RIR 184
Db 179 SIR 181
RESULT 14
Q945S2 PRELIMINARY; PRT; 563 AA.
AC Q945S2
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2003 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA
DE N-glycosidase).
GN ABL.
OS *Sambucus nigra* (European elder).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; campanulids; Dipsacales; Adoxaceae; Sambucus.

```

QY      7 YPIINFNTAGATVQSYNTFNIRAVR-----GRLTVLPNRVGLPIINORFTLVLVLSNH 56
Db      28 YPSVSENLGDGAKSATYRDFLSNRKKTVACTYEVNGLPVLIRRESEVOVKSRFVLVPLTNY 87
QY      57 AELSVTIALDVNAYVGYVYAGNSATFFHPDNOEDABAI--THLPTDVQNRYYTFAFGNY 114
Db      88 NGNVITVLAVDVNIVYVAFSGNANSYFF-----KDATEVQKSNLFGYTKQN-TLSEFTGY 141
QY      115 DRLEQLAGNLFRENTLNGNGLEPAISALYYSTGGTQLPTLARSFFICIMISEAARFOY 174
Db      142 DNLETAANTRRSEIELGPSPLDGAITSLYHGD-----SVARSLLVVIQWYSEARFXY 194
QY      175 IEGEMRTIR 184
Db      195 IEQVRRSLQ 204

```

Search completed: February 10, 2004, 16:26:37
Job time : 25:2187 secs

Search completed: February 10, 2004, 16:26:37
Job time : 25.2187 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:13:20 ; Search time 33.7583 Seconds
(without alignments)
930.966 Million cell updates/sec

Title: US-10-083-336A-7

Perfect score: 1019

Sequence: 1 MVPKQYPIINFAGTAGTQVS.....ARFQVIEGEMRTRINRRS 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_15Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1019	100.0	267 13	AAR30722 Ricin A from pIC11
2	1019	100.0	267 21	AAB19265 Amino acid sequenc
3	1010	99.1	267 14	AAR37290 Ricin A chain. Un
4	1010	99.1	267 16	AAR63902 Ricin A-chain (RTA
5	1010	99.1	290 18	AAW25136 Ricin A-chain ribo
6	1010	99.1	290 18	AAW21699 Ricin A-chain RIP.
7	1010	99.1	332 8	AAP70097 Ricin A. Escheric
8	1010	99.1	332 8	AAP70838 Sequence of Ricinu
9	1010	99.1	332 10	AAP95639 Ricin A encoded by

10	1010	99.1	554	16	AAR70827	Anti-cataract immu
11	1010	99.1	562	10	AAP90079	Ricin D. Ricinus
12	1010	99.1	565	22	AAG78304	Modified castor be
13	1010	99.1	576	8	AAP70326	Sequence of Ricinu
14	1010	99.1	576	18	AAW25787	Castorbean ricin.
15	1010	99.1	576	20	AAW55892	Castor bean ricin
16	1010	99.1	576	21	AAW78592	Ricinus communis r
17	1010	99.1	576	22	AAG78301	Castor bean prepro
18	1010	99.1	576	22	AAG78302	Castor bean prepro
19	1007	98.8	565	6	AAP50166	Sequence of prepro
20	1007	98.8	565	22	AAG78300	Castor bean prepro
21	1006	98.7	200	9	AAP80164	Biosynthetic multi
22	1003	98.4	268	14	AAR39570	Sequence of ricin-
23	1003	98.4	574	8	AAP70325	Sequence of Ricinu
24	1002	98.3	565	7	AAP60240	preproricin. Rici
25	1001	98.2	574	10	AAP94793	DNA sequence of ri
26	1000	98.1	534	14	AAR39571	Sequence of G-FIT.
27	999	98.0	332	11	AAR06554	Ricin A gene produ
28	996	97.7	267	14	AAR32430	Ricin A. Syntheti
29	983	96.5	267	16	AAW74176	Ricin A chain (RTA
30	920.5	90.3	540	18	AAW25143	Castor oil plant a
31	920.5	90.3	540	18	AAW21706	R. communis agglut
32	786.5	77.2	534	8	AAP70324	Sequence of Ricin
33	778	76.3	280	10	AAP95648	Ricin agglutinin A
34	342	33.6	247	16	AAR67359	Trichosanthin anti
35	342	33.6	247	21	AAW69048	Amino acid sequenc
36	342	33.6	248	11	AAW07518	Synthetic alpha-tr
37	342	33.6	248	13	AAW25573	Mature alpha-Trich
38	342	33.6	267	18	AAW25140	Trichosanthin (a r
39	342	33.6	267	18	AAW21703	Trichosanthin. Tr
40	342	33.6	289	11	AAW07514	Trichosanthin from
41	342	33.6	289	13	AAW25572	Trichosanthin prot
42	342	33.6	289	14	AAW32986	Encodes chinese cu
43	342	33.6	289	15	AAW55129	Alpha-trichosanthi
44	342	33.6	289	15	AAW55129	Chinese cucumber a
45	342	33.6	289	18	AAW10468	

ALIGNMENTS

RESULT 1

AAR30722
ID AAR30722 standard; Protein; 267 AA.
XX AAR30722;
XX AC
XX DT 25-MAR-2003 (updated)
XX DT 08-FEB-1993 (first entry)
XX DE Ricin A from pIC1102.
XX KW pH; temperature; cultivation; host; soluble.
XX OS Synthetic.
XX PN EP501692-A2.
XX PD 02-SEP-1992.
XX PF 21-FEB-1992; 92EP-0301466.
XX PR 26-FEB-1991; 91GB-0003925.
XX PR 26-FEB-1991; 91GB-0003926.
XX PR 26-FEB-1991; 91GB-0004016.
XX (ICIL) IMPERIAL CHEM IND PLC.
XX (ZENE) ZENECA LTD.
XX Fitton JE, Hockney RC, Kara BV;
XX WPI; 1992-294124/36.
XX N-PSDB; RAQ27876.

XX Prepn. of soluble recombinant polypeptide(s), esp. ricin A - by
 PT adjusting pH and/or temp. during cultivation to increase yield of
 XX soluble protein.
 XX Disclosure; Fig 9; 49pp; English.
 XX
 XX Prepn. of ricin A comprises cultivating a host, including a DNA
 CC sequence which encodes ricin A, e.g. from pCil102, in a nutrient
 CC medium for an initial period at a first pH value which favours growth
 CC of the host; and cultivation of the host for a further period at a pH
 CC lower than the first pH value, and opt. cooling the host during the
 CC terminal portion of the cultivation and harvesting the host during
 CC the terminal portion. By adjusting the pH and temp. during the
 CC cultivation of hosts, high yields of soluble recombinant ricin A
 CC are obtained.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX
 XX SQ Sequence 267 AA;

Query Match 100.0%; Score 1019; DB 13; Length 267;
 Best Local Similarity 100.0%; Pred. No. 5.8e-100;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRGLPINQRFILV 60
 Db 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRGLPINQRFILV 60
 Qy 61 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 120
 Db 61 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 120
 Qy 121 GNYDRLEQLAGNLRENIELNGPLEEASALYYSTGGTQPTLARSFIICQMISEAAR 180
 Db 121 GNYDRLEQLAGNLRENIELNGPLEEASALYYSTGGTQPTLARSFIICQMISEAAR 180
 Qy 181 FOYIEGEMTRIRYNRRS 198
 Db 181 FOYIEGEMTRIRYNRRS 198

RESULT 2
 AAB19265
 ID AAB19265 standard; protein; 267 AA.
 XX
 AC AAB19265;
 XX
 DT 19-FEB-2001 (first entry)
 XX
 DE Amino acid sequence of a human ricin toxin A chain (RTA).
 XX
 KW Immunotoxin; cytokine; vascular leak syndrome; VLS; lymphoma; myeloma;
 KW Graft versus host disease; metastatic lesion tumour; tumour; immunotoxin;
 KW ricin toxin A chain.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 74..76
 FT /note= "vascular leak syndrome (VLS) inducing motif"
 XX
 FN WO200058456-A2.
 XX
 PD 05-OCT-2000.
 XX
 PF 30-MAR-2000; 2000WO-US08600.
 XX
 PR 30-MAR-1999; 99US-0126826.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 XX
 PI Vitetta ES, Ghetie VF, Baluna RG, Smallshaw J;

XX WPI; 2000-664922/64.
 XX
 XX Modifying the ability of a proteinaceous composition to induce a toxic
 PT effect for reducing vascular leak syndrome, comprises identifying at
 PT least one specified amino acid sequence and altering it -
 XX
 XX Example 1; Page 119-120; 125pp; English.
 XX
 CC The specification describes a method for producing immunotoxins and
 CC cytokines with a reduced ability to promote vascular leak syndrome
 CC (VLS). The immunotoxins are useful for treating graft versus host
 CC disease, non-Hodgkin's and Hodgkin's lymphoma, myeloma, metastatic
 CC lesion tumours and some type of solid tumours. The present sequence
 CC represents an immunotoxin (ricin toxin A chain) which induces VLS.
 CC The VLS-inducing motif can be mutated or deleted so that VLS is not
 CC induced.
 XX
 XX SQ Sequence 267 AA;

Query Match 100.0%; Score 1019; DB 21; Length 267;
 Best Local Similarity 100.0%; Pred. No. 5.8e-100;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRGLPINQRFILV 60
 Db 1 MVPKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRGLPINQRFILV 60
 Qy 61 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 120
 Db 61 ELSNHAELSVTLALDVNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYTFAFG 120
 Qy 121 GNYDRLEQLAGNLRENIELNGPLEEASALYYSTGGTQPTLARSFIICQMISEAAR 180
 Db 121 GNYDRLEQLAGNLRENIELNGPLEEASALYYSTGGTQPTLARSFIICQMISEAAR 180
 Qy 181 FOYIEGEMTRIRYNRRS 198
 Db 181 FOYIEGEMTRIRYNRRS 198

RESULT 3
 AAR37290
 ID AAR37290 standard; protein; 267 AA.
 XX
 AC AAR37290;
 XX
 DT 25-MAR-2003 (updated)
 DT 09-JAN-2003 (updated)
 DT 13-SEP-1993 (first entry)
 XX
 DE Ricin A chain.
 XX
 KW Type II ribosome-inactivating protein; type II RIP; gelonin;
 KW momordin; immunoconjugate; autoimmune disease; cell killing; toxin.
 XX
 OS Unidentified.
 XX
 FN WO9309130-A1.
 XX
 PD 13-MAY-1993.
 XX
 PF 04-NOV-1992; 92WO-US09487.
 XX
 PR 04-NOV-1991; 91US-0787567.
 PR 19-JUN-1992; 92US-0901707.
 XX
 PA (XOMA) XOMA CORP.
 XX
 PI Berhard SL, Better MD, Carroll SF, Lane JA, Lei SP;
 XX WPI; 1993-167617/20.
 XX

PT Analogues of type I ribosome inactivating protein - useful as
 PT cytotoxic agents, immuno toxins for treating auto immune diseases,
 PT cancer, graft versus host disease and selective cell killing in-vivo
 XX
 PS Claim 1; Page 92; 163pp; English.

XX The invention covers analogues of Type I RIPs. Ricin is a Type II
 CC RIP whose A chain is homologous to plant type I RIPs. The analogues
 CC of the invention have a cysteine available for intermolecular
 CC disulphide bonding at an amino acid position corresp. to a position
 CC not naturally available for bonding; the cys residue is located in
 CC the C-terminal region of the analogue between a position corresp. to
 CC amino acid 251 and the C-terminus of ricin A chain. The analogues are
 CC pref. joined via a disulphide linkage to a molecule which specifically
 CC binds to a target cell, e.g. an antibody fragment.
 CC (Updated on 09-JAN-2003 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 267 AA;
 Query Match 99.1%; Score 1010; DB 14; Length 267;
 Best Local Similarity 100.0%; Pred. No. 5.2e-99;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVPLPNRVGLPINQRFILVEL 62
 Db 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVPLPNRVGLPINQRFILVEL 62
 QY 63 SNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFGN 122
 Db 63 SNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFGN 122
 QY 123 YDRLEQLAGNLRNLELNGPLERALSALYYSTGGTOLPTLARSFIIQIMISEAARFQ 182
 Db 123 YDRLEQLAGNLRNLELNGPLERALSALYYSTGGTOLPTLARSFIIQIMISEAARFQ 182
 QY 183 YIEGEMRTRIRYNRRS 198
 Db 183 YIEGEMRTRIRYNRRS 198

RESULT 4
 AAR63902
 ID AAR63902 standard; protein; 267 AA.
 AC AAR63902;
 XX

DT 25-MAR-2003 (updated)
 DT 27-JUL-1995 (first entry)
 XX

DE Ricin A-chain (RTA).

XX Ricin A chain; RTA; ribosome-inactivating proteins; RIPs;
 XX cytotoxic therapeutic agents; autoimmune disease; cancer;
 XX graft-versus-host disease.

XX Ricinus communis.

XX W09426910-A1.

XX 24-NOV-1994.

XX 12-MAY-1994; 94WO-US05348.

XX 12-MAY-1993; 93US-0064691.

XX (XOMA) XOMA CORP.

XX Better MD, Carroll SS, Studnicka GM, Carroll SF;
 PI WPI; 1995-006804/01.

XX Polynucleotide(s) encoding type I ribosome-inactivating proteins

PT

PT - which are suitable for use as components of cytotoxic
 PT therapeutic agents.

XX Example 3; Fig 1; 221pp; English.

XX AAR63902 is the ricin A chain gene product, it is analogous to the
 CC ribosome-inactivating proteins (RIPs) described in AAR63903-R63911.
 CC RIPs are the key components of cytotoxic therapeutic agents (CTAs),
 CC which include gene fusion products and immunoconjugates. CTAs may
 CC be used to selectively eliminate any cell type to which a RIP
 CC component is targeted, by the specific binding capacity of the
 CC second component of the agent. They can be used in the treatment
 CC of diseases where the elimination of a particular cell type is
 CC desired, such as autoimmune disease, cancer and graft-versus-host
 CC disease.
 CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 267 AA;

Query Match 99.1%; Score 1010; DB 16; Length 267;
 Best Local Similarity 100.0%; Pred. No. 5.2e-99;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVPLPNRVGLPINQRFILVEL 62
 Db 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHIVPLPNRVGLPINQRFILVEL 62
 QY 63 SNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFGN 122
 Db 63 SNHAELSVTLALDVNTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNYRTFAFGN 122
 QY 123 YDRLEQLAGNLRNLELNGPLERALSALYYSTGGTOLPTLARSFIIQIMISEAARFQ 182
 Db 123 YDRLEQLAGNLRNLELNGPLERALSALYYSTGGTOLPTLARSFIIQIMISEAARFQ 182
 QY 183 YIEGEMRTRIRYNRRS 198
 Db 183 YIEGEMRTRIRYNRRS 198

RESULT 5
 AAW25136
 ID AAW25136 standard; Protein; 290 AA.
 AC AAW25136;
 XX

DT 25-MAR-2003 (updated)
 DT 02-DEC-1997 (first entry)
 XX

DE Ricin A-chain ribosome inhibitory protein inactive precursor.

XX Maize; proRIP; ribosome inactivating protein; alpha; beta subunit;
 XX internal linker; Barley Translation Inhibitor; Trichosanthin;
 XX Ricin A-chain; Abrin-A A-chain; Saporin; SUT-1; Luffin A; WAP;
 XX Ricinus communis agglutinin; Momordin; PAP-S; Luffin-B; Dianthin 30;
 XX therapeutic toxin; tumour cell targeted; protein synthesis inhibitor;
 XX post-translational modification; cancer; neoplasia; HIV; AIDS;
 XX human immunodeficiency virus; acquired immune deficiency syndrome.

XX Synthetic.

XX US5646026-A.

XX 08-JUL-1997.

XX 07-JUN-1995; 95US-0485286.

XX 09-DEC-1992; 92US-0987927.

XX 11-JUN-1990; 90US-0535636.

XX 26-JAN-1995; 95US-0378761.

XX 07-JUN-1995; 95US-0485286.

XX (DOWC) DOWELANCO.

XX	Hey TD, Morgan AER, Walsh TA;
PI	WPI; 1997-362934/33.
XX	
DR	DNA encoding pro-ribosome inactivating proteins - inactive
PT	precursors of ribosome inactivating proteins; can be expressed in
PT	eukaryotic cells without causing cell death
XX	
PS	Claim 4; Column 91-94; 186pp; English.
XX	
CC	AAW25136 represents a Ricin A-chain ribosome inhibitory protein (RIP)
CC	which was engineered to contain a selectively removable internal peptide
CC	linker sequence separating the alpha and beta units of the RIP. When
CC	separated the two units regain activity and are capable of inactivating
CC	eukaryotic ribosomes and hence preventing protein production. Many
CC	different RIPs may be produced with an internal linker including
CC	maize RIP, Trichosanthin, Ricin A-chain, Abrin-A A-chain and
CC	Saporin. The RIPs can be used in the construction of therapeutic
CC	toxins targeted to specific cells such as tumour cells via the
CC	attachment of a targeting polypeptide, e.g. a monoclonal antibody.
CC	A further use is in HIV therapy (see US4869903). There is interest
CC	in expressing RIP recombinantly in host eukaryotic cells, because of
CC	the capacity to provide correct post-translational processing. However,
CC	RIPs effectively inhibit protein synthesis in eukaryotic cells resulting
CC	in cell death. Since the inactive RIP proteins are not cytotoxic to
CC	eukaryotic cells, they can be recombinantly expressed in such cells and
CC	then converted to active RIP proteins.
CC	(Updated on 25-MAR-2003 to correct PF field.)
XX	
SQ	Sequence 290 AA;
	Query Match 99.1%; Score 1010; DB 18; Length 290;
	Best Local Similarity 100.0%; Pred. No. 5.8e-99;
	Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Qy	3 PKQYPIINFTHAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRVGLPINQRFILVEL 62
Db	27 PKQYPIINFTHAGATVQSYTNFIRAVRGRLTTGADVREHIVLPNVRVGLPINQRFILVEL 86
Qy	63 SNHAELSVTLALDVTNAVYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNYRTAFGNG 122
Db	87 SNHAELSVTLALDVTNAVYVGYRAGNSAYFFHPDQEDAEAIHTLFTDVQNYRTAFGNG 146
Qy	123 YDRLEQLAGNLRENIELNGNGPLEEASISALYYVYSTGGTOLPTLARSFIICIMISEARFQ 182
Db	147 YDRLEQLAGNLRENIELNGNGPLEEASISALYYVYSTGGTOLPTLARSFIICIMISEARFQ 206
Qy	183 YIEGEMETRIYNNRS 198
Db	207 YIEGEMETRIYNNRS 222
RESULT 6	
AAW21699	
ID	AAW21699 standard; Protein; 290 AA.
XX	
AC	AAW21699;
XX	
DT	25-MAR-2003 (updated)
DT	26-SEP-1997 (first entry)
XX	
DE	Ricin A-chain RIP.
XX	
KW	pro-Ribosome Inactivating Protein; proRIP; peptide linker; cancer;
KW	inactivation; eukaryotic ribosome; alpha fragment; beta fragment;
KW	inhibitor; protein synthesis; N-glycosidase; glycosidic bond; liver;
KW	rat; ribosomal 28S RNA; cellular proliferation; HIV-infected T cell.
OS	Ricinus communis.
XX	
Key	Location/Qualifiers
FT	Region 152..162

```

XX Ricin A; Met-aminopeptidase.
XX KW Escherichia coli.
XX OS
XX EP219237-A.
XX PD 22-APR-1987.
XX PF 19-SEP-1986; 86EP-0307242.
XX PR 06-MAY-1986; 86US-0860330.
XX PR 20-SEP-1985; 85US-0778414.
XX PA (CETU ) CETUS CORP.
XX PI Benbassat A, Bauer KA, Chang S, Chang SY;
XX WPI; 1987-110172/16.
XX DR N-PSDB; AAN70152.
XX PT N-terminal methionine free proteins prodn. - by using host
XX PT transformed with vector to express a methionine-amino-peptidase
XX PS Disclosure; Fig. 4; 20pp; English.
XX CC Ricin A may be produced in a form which lacks an N-terminal Met
XX CC using Met-aminopeptidase from E.coli.
XX SQ Sequence 332 AA;

Query Match          99.1%; Score 1010; DB 8; Length 332;
Best Local Similarity 100.0%; Pred. No. 7e-99;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGLPINQRFILVEL 62
DQ 38 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGLPINQRFILVEL 97
QY 63 SNHAELSVTLALDVNTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTEAFGNG 122
DQ 98 SNHAELSVTLALDVNTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTEAFGNG 157
QY 123 YDRLEQLAGNLRNIELNGPLEEAISALYYSTGGTQPLTARSFFIICMISEARFQ 182
DQ 158 YDRLEQLAGNLRNIELNGPLEEAISALYYSTGGTQPLTARSFFIICMISEARFQ 217
QY 183 YIEGEMTRIRYNRRS 198
DQ 218 YIEGEMTRIRYNRRS 233

RESULT 8
AAP70838
ID AAP70838 standard; protein; 332 AA.
XX AC AAP70838;
XX AC
XX DT 25-MAR-2003 (updated)
XX DT 18-FEB-1991 (first entry)
XX DE Sequence of Ricinus communis castor beans ricin toxin (RT or ricin)
XX DE A protein encoded by PRA123.
XX KW Lectin; toxin protein; cytotoxic; castor bean;
XX KW plant toxin.
XX OS Ricinus communis.
XX FH Key Location/Qualifiers
XX FT Region 1..332
XX FT /note="leader"
XX FT Region 33..302

```

```

FT FT /note="A-chain"
FT FT 315..332
FT FT /note="B-chain"
XX PN EP237676-A.
XX PD 23-SEP-1987.
XX PF 13-NOV-1986; 86EP-0308877.
XX PR 07-MAR-1986; 86US-0837583.
XX PA (CETU ) CETUS CORP.
XX PA (CHIR ) CHIRON CORP.
XX PI Piatak M;
XX WPI; 1987-265177/38.
XX DR N-PSDB; AAN70519.
XX PT New non-glycosylated ricin precursor and toxin etc. - are prepd.
XX PT by recombinant DNA procedures with specific isolation steps for
XX PT purer and soluble prods.
XX PS Disclosure; Fig 1; 112pp; English.
XX CC The full-length sequences encoding ricin A (AAN70520), ricin D
XX CC (AAN70525), putative ricin E (AAN70526) and RCA (AAN70524) in precursor
XX CC form were obtd. using messenger RNA to obtain a cDNA library, and
XX CC then probing the library to retrieve the desired cDNA inserts. The
XX CC library was probed using the 35-mer given in AAN70514. Figure 4 (see
XX CC AAN70520, AAN70521, AAN70522) shows the nucleotide sequences of three
XX CC plasmids contg. cDNA inserts obtd. by probing a cDNA library for
XX CC sequences encoding ricin B using the probe in AAN70517. The cDNA
XX CC inserts can be placed into expression vectors. Site-directed
XX CC mutagenesis may be used to place an ATG start codon and a HindIII
XX CC site at the beginning of the mature protein (see AAN70518). The
XX CC coding sequences of the inserts can be ligated into expression
XX CC vectors contg. the phoA promoter-operator and leader sequence
XX CC (AAN70523) and suitable retroregulators.
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 332 AA;

Query Match          99.1%; Score 1010; DB 8; Length 332;
Best Local Similarity 100.0%; Pred. No. 7e-99;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGLPINQRFILVEL 62
DQ 38 PKQYPIINFTTAGATVQSYTNFIRAVRGRLTTGADVREHPIVLPNVRGLPINQRFILVEL 97
QY 63 SNHAELSVTLALDVNTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTEAFGNG 122
DQ 98 SNHAELSVTLALDVNTNAYVGVYRAGNSAYFFHPDNOEDAEATHLFTDVQNYRTEAFGNG 157
QY 123 YDRLEQLAGNLRNIELNGPLEEAISALYYSTGGTQPLTARSFFIICMISEARFQ 182
DQ 158 YDRLEQLAGNLRNIELNGPLEEAISALYYSTGGTQPLTARSFFIICMISEARFQ 217
QY 183 YIEGEMTRIRYNRRS 198
DQ 218 YIEGEMTRIRYNRRS 233

RESULT 9
AAP95639
ID AAP95639 standard; protein; 332 AA.
XX AC AAP95639;
XX AC
XX DT 25-MAR-2003 (updated)
XX DT 31-OCT-2002 (updated)

```

DT 13-AUG-1990 (first entry)
 XX Ricin A encoded by insert from plasmid pRA123.
 DE Plasmid pRA123; ricin-A; ricin-B; cytotoxicity.
 XX Ricinus communis.
 OS Synthetic.
 XX Key Location/Qualifiers
 FH 1..35
 FT Peptide /label= leader sequence
 FT Peptide 36..302
 FT Peptide /label=A-chain
 FT Peptide 303..314
 FT Peptide /label=linker
 FT Peptide 315..332
 FT Peptide /label=B-chain
 XX EP335476-A.
 XX 04-OCT-1989.
 XX 19-JAN-1989; 89EP-0201162.
 XX 08-FEB-1984; 84US-0578115.
 XX 08-FEB-1984; 84US-0578121.
 XX 09-FEB-1984; 84US-0578122.
 XX 07-SEP-1984; 84US-0648759.
 XX 20-SEP-1984; 84US-0653515.
 XX (CETU) CETUS CORPORATION.
 XX Gelfand D, Lawyer FC, Horn G, Greenfield L, Nitecki D, Kaplan D;
 FI Platak MJ;
 XX WPI; 1989-286959/40.
 DR N-PSDB; AAN91281.
 XX Recombinant vectors expressing ricin chains or diphtheria toxin -used for
 PT prodn. of new immunotoxin conjugates with monoclonal antibodies, having
 FT high cell specificity and good extracellular stability.
 XX Disclosure; Fig 14; 54pp; English.
 XX Plasmid RA123 (ATCC No. 39799) carries the entire coding sequence for
 CC ricin A, as well as codons for 12 AAs joining the A to the B chain.
 CC Following modification for ease of manipulation the plasmid was used to
 CC construct expression vectors which express the conjugates in
 CC host cells.
 CC (Updated on 31-OCT-2002 to add missing OS field.)
 CC (Updated on 25-MAR-2003 to correct PF field.)
 CC (Updated on 25-MAR-2003 to correct PR field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX Sequence 332 AA;
 SQ Query Match 99.1%; Score 1010; DB 10; Length 332;
 Best Local Similarity 100.0%; Pred. No. 7e-99;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3 PKQYPIINFATTAGATVQSYNFINRAVRGRLLTGADVREHPVLPNVRVGLPINQRFILVEL 62
 DB 38 PKQYPIINFATTAGATVQSYNFINRAVRGRLLTGADVREHPVLPNVRVGLPINQRFILVEL 97
 OY 63 SNHAELSVTALDVNTNAYVGVYAGNSAYFPHPDQDAEAITHLFTDVQNYRTFAFGGN 122
 DB 98 SNHAELSVTALDVNTNAYVGVYAGNSAYFPHPDQDAEAITHLFTDVQNYRTFAFGGN 157
 OY 123 YDRLEQAGNLRENIEFGNGLEPAISALYYSTGGTQLPTLARSFICQMISEAARFQ 182
 DB 158 YDRLEQAGNLRENIEFGNGLEPAISALYYSTGGTQLPTLARSFICQMISEAARFQ 217

OY 183 YIEGEMTRIRYNNRS 198
 DB 218 YIEGEMTRIRYNNRS 233
 RESULT 10
 AAR70827
 ID AAR70827 standard; Protein; 554 AA.
 XX AAR70827;
 AC
 XX 25-MAR-2003 (updated)
 DT 31-AUG-1995 (first entry)
 XX Anti-cataract immunotoxin.
 XX Immunotoxin; heavy chain; light chain; variable region; antibody;
 KW ricin-A; cytostatic; cataract; lens opacification; epithelial cell;
 KW pHB19; 4197X; monoclonal antibody; MAb.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH 1..27
 FT Peptide /label= sig_peptide
 FT /note= "phoA signal sequence"
 FT Domain 28..145
 FT /label= HEAVY
 FT /note= "Mab 4197X heavy chain"
 FT Peptide 148..166
 FT /label= LINKER
 FT Domain 169..274
 FT /label= LIGHT
 FT /note= "Mab 419X light chain"
 FT Domain 276..544
 FT /label= RICIN-A
 FT Peptide 549..554
 FT /label= TAG
 FT /note= "hexa-histidine tail"
 XX WO9503828-A1.
 XX 09-FEB-1995.
 XX 15-JUL-1994; 94WO-US07919.
 XX 02-AUG-1993; 93US-0101329.
 XX (HOUS-) HOUSTON BIOTECHNOLOGY INC.
 XX Gould RM, Kelleher PJ, Wallace TL, Wood MS;
 XX WPI; 1995-082036/11.
 DR N-PSDB; AAQ85386.
 XX New single chain immuno:toxin - binds specifically to epithelial
 PT cells, for inhibiting development of sec. cataracts after
 PT extra:capular cataract extraction.
 XX Disclosure; Fig.4; 68pp; English.
 XX The immunotoxin given in AAR70827 comprises the heavy and light chain
 CC variable regions of anti-lens epithelium IGG3 MAb 4197X linked to
 CC ricin-A and a hexa-histidine tag. The DNA construct encoding the
 CC immunotoxin was expressed from pHB19 in E. coli.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX Sequence 554 AA;
 SQ Query Match 99.1%; Score 1010; DB 16; Length 554;
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVEL 62
 Db 280 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVEL 339
 QY 63 SNHAELSVTALDVTNAYVVGYSAGNSAYFFHPDNOQEDAEATHLFTDQNRVYTFAGGN 122
 Db 340 SNHAELSVTALDVTNAYVVGYSAGNSAYFFHPDNOQEDAEATHLFTDQNRVYTFAGGN 399
 QY 123 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQLTPLARSFIIICQMISEARFQ 182
 Db 400 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQLTPLARSFIIICQMISEARFQ 459
 QY 183 YIEGEMTRIRYNRRS 198
 Db 460 YIEGEMTRIRYNRRS 475

RESULT 11
 AAP90079
 ID AAP90079 standard; protein; 562 AA.
 AC
 XX AAP90079;
 XX
 DT 25-MAR-2003 (updated)
 DT 01-NOV-1989 (first entry)
 XX
 DE Ricin D.
 XX
 KW Ricin D; Ricinus communis; castor beans; Zambiaricinis variety;
 KW modified; lectin binding removed; reduced cell binding
 XX
 OS Ricinus communis (castor beans).
 XX
 PN WO8904839-A.
 XX
 PD 01-JUN-1989.
 XX
 PF 23-NOV-1988; 88WO-US04238.
 XX
 PR 24-NOV-1987; 87US-0124735.
 XX
 PA (GEMY) GENETICS INST INC.
 XX
 PI Brown EL, Jones S;
 XX
 DR WPI; 1989-178366/24.
 XX
 DR N-PSDB; AAN90068.
 XX
 XX Modified ricin molecules and toxin conjugates
 PT - in which the lectin binding function of the B chain
 PT is removed or diminished to reduce cell binding.
 XX
 PS Disclosure; fig 1; 51pp; English.
 XX
 CC Ricin D (see corresp. AAN90068) found on the EcoRI-HindIII fragment
 CC of DNA from Ricinus communis, Zambiaricinis variety. Patent
 CC discloses many modifications of ricin in which the lectin binding
 CC function of the B chain is diminished or removed, and conjugation
 CC to toxins to eliminate cell binding.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 562 AA;

Query Match 99.1%; Score 1010; DB 10; Length 562;
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVEL 62
 Db 38 PKQYPIINFTTAGATVQSYTNFIRAVRGLTTGADVRHEIPVLPNRVGLPINQRFILVEL 97
 QY 63 SNHAELSVTALDVTNAYVVGYSAGNSAYFFHPDNOQEDAEATHLFTDQNRVYTFAGGN 122

Db 98 SNHAELSVTALDVTNAYVVGYSAGNSAYFFHPDNOQEDAEATHLFTDQNRVYTFAGGN 157
 QY 123 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQLTPLARSFIIICQMISEARFQ 182
 Db 158 YDRLEQLAGNLRENIELNGPLEEALISALYYSTGGTQLTPLARSFIIICQMISEARFQ 217
 QY 183 YIEGEMTRIRYNRRS 198
 Db 218 YIEGEMTRIRYNRRS 233

RESULT 12
 AAG78304
 ID AAG78304 standard; Protein; 565 AA.
 XX
 AC AAG78304;
 XX
 DT 27-NOV-2001 (first entry)
 XX
 DE Modified castor bean preprorizin (SEQ ID 10).
 XX
 KW Castor bean plant; preprorizin; ricin; A chain; B chain;
 KW human immunodeficiency virus infection; HIV; toxin; antiviral agent;
 KW retroviral infection; anti-HIV; virucide; viral protease.
 XX
 OS Chimeric - Ricinus communis
 OS Chimeric - Human immunodeficiency virus type 2.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..24
 FT /label= Signal peptide
 FT 25..565
 FT /label= Prorizin
 FT /note= "Prorizin consists of the ricin A chain, a linker
 FT peptide, and the ricin B chain. Prorizin is
 FT proteolytically cleaved between the A chain and
 FT the linker to yield mature ricin"
 XX
 FT Protein 25..291
 FT /label= Ricin A chain
 FT /note= "N-glycosidase"
 FT 292..303
 FT /label= Linker peptide
 FT 296..297
 FT /label= HIV protease cleavage site
 FT 304..565
 FT /label= Ricin B chain
 FT /note= "Galactose/N-acetylglactosamine-binding lectin"
 XX
 WO200160393-A1.
 XX
 PD 23-AUG-2001.
 XX
 PF 15-FEB-2001; 2001WO-US05282.
 XX
 PR 16-FEB-2000; 2000US-0182759.
 XX
 PA (BECH-) BECHTEL BWXT IDAHO LLC.
 XX
 PI Keener WK, Ward TE;
 XX
 DR WPI; 2001-581908/65.
 DR N-PSDB; AAI64145.
 XX
 XX Novel composition comprising toxin e.g., ricin based antiviral compound
 PT useful for treating viral infections such as human immunodeficiency
 PT virus infection.
 XX
 PS Example 1; Page 59-63; 66pp; English.
 XX
 CC The sequence relates to the amino acid sequence of a modified preprorizin
 CC protein encoded by AAI64145. The invention relates to a novel toxin
 CC (e.g. ricin) based antiviral agent which is toxic to virus-infected

CC cells, but non-toxic to uninfected cells. The invention has anti-HIV and
 CC virucide activities. The agent is able to enter all HIV susceptible
 CC cells, and not just cells known to act as host cells for the virus. The
 CC antiviral agent remains inert in a cell unless the cell is infected
 CC with the HIV virus, where the viral protease activates it. Ricin's
 CC mechanism of action is through inactivation of cellular ribosomes and
 CC enhancement of binding of the antiviral agent to galactose residues on
 CC cell surfaces, and its cellular internalisation. The invention is useful
 CC for treating human immunodeficiency virus infection and other viral
 CC infections, especially retroviral infections. The antiviral agent is
 CC activated in viral particles or early-stage infected cells, killing the
 CC cells upon infection and effectively preventing the integration of the
 CC viral genome into the host genome thereby preventing the latency/rebound
 CC problem.

XX SQ Sequence 565 AA;
 Query Match 99.1%; Score 1010; DB 22; Length 565;
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PKQYPIINFNTAGATVQSYTNFIRAVRGLTTGADVREHPVLPNRVGLPINQRFILVEL 62
 DB 27 PKQYPIINFNTAGATVQSYTNFIRAVRGLTTGADVREHPVLPNRVGLPINQRFILVEL 86
 QY 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGN 122
 DB 87 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGN 146
 QY 123 YDRLEQLAGNLRNIELGNGPLEEASALYYSTGTQTLPTLARSFFICIQMISEAARFQ 182
 DB 147 YDRLEQLAGNLRNIELGNGPLEEASALYYSTGTQTLPTLARSFFICIQMISEAARFQ 206
 QY 183 YIEGEMTRIRYNRRS 198
 DB 207 YIEGEMTRIRYNRRS 222

RESULT 13
 AAP70326
 ID AAP70326 standard; Protein; 576 AA.
 AC AAP70326;
 XX
 XX Key Location/Qualifiers
 XX Region 1..35
 FT FT /note= "leader"
 FT FT 36..302
 FT FT /note= "A-chain"
 FT FT 315..576
 FT FT /note= "B-chain"
 XX EP237676-A.
 PN
 XX
 PD 23-SEP-1987.
 XX
 XX 13-NOV-1986; 86EP-0308877.
 PF
 XX
 XX 07-MAR-1986; 86US-0837583.
 PR
 XX
 XX (CETU) CETUS CORP.
 PA
 PA (CHIR) CHIRON CORP.

Lectin; toxin protein; cytotoxic; cytostatic; castor bean;
 plant toxin.

Ricinus communis.

XX PIatak M;
 FI
 XX WPI; 1987-265177/38.
 DR N-PSDB; AAN70526.
 DR

XX New non-glycosylated ricin precursor and toxin etc. - are prepd.
 FT by recombinant DNA procedures with specific isolation steps for
 FT purer and soluble prods.
 PS

XX Disclosure; Fig 14(1-2); 112pp; English.

XX The full length sequences encoding ricin A (AAN70520), ricin D
 CC (AAN70525) putative ricin E (AAN70526) and RCA (AAN70524) in precursor
 CC form were obtained, using the messenger RNA to obtain a cDNA library, and
 CC then probing the library to retrieve the desired cDNA inserts. The
 CC library was probed using the 35-mer given in AAN70514. Figure 4 (see
 CC AAN70520, AAN70521, AAN70522), shows the nucleotide sequences of three
 CC plasmids containing cDNA inserts obtained by probing a cDNA library
 CC for sequences encoding ricin B using the probe in AAN70517. The cDNA
 CC inserts can be placed into expression vectors. Site-directed
 CC mutagenesis may be used to place an ATG start codon and a HindIII
 CC site at the beginning of the mature protein, (see AAN70518). The
 CC coding sequences of the inserts can be ligated into expression
 CC vectors containing the Phoa promoter-operator and leader sequence
 CC (AAN70523) and suitable retroregulators.
 CC (Updated on 25-MAR-2003 to correct PA field.)
 XX

SQ Sequence 576 AA;

Query Match 99.1%; Score 1010; DB 8; Length 576;
 Best Local Similarity 100.0%; Pred. No. 1.5e-98;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFNTAGATVQSYTNFIRAVRGLTTGADVREHPVLPNRVGLPINQRFILVEL 62
 DB 38 PKQYPIINFNTAGATVQSYTNFIRAVRGLTTGADVREHPVLPNRVGLPINQRFILVEL 97
 QY 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGN 122
 DB 98 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEATHLFTDVQNRYYTFAFGN 157
 QY 123 YDRLEQLAGNLRNIELGNGPLEEASALYYSTGTQTLPTLARSFFICIQMISEAARFQ 182
 DB 158 YDRLEQLAGNLRNIELGNGPLEEASALYYSTGTQTLPTLARSFFICIQMISEAARFQ 217
 QY 183 YIEGEMTRIRYNRRS 198
 DB 218 YIEGEMTRIRYNRRS 233

RESULT 14

AAN25787
 ID AAN25787 standard; Protein; 576 AA.

AC AAN25787;
 XX
 DT 25-MAR-2003 (updated)
 DT 27-MAR-1998 (first entry)
 XX
 DE Castorbean ricin.

Ricin; cytotoxin; hybrid protein; cell delivery;
 cell binding ligand; translocation domain; diphtheria toxin B';
 interleukin-2; T-cell lymphoma; organ rejection; therapy.

Ricinus communis.

XX Key Location/Qualifiers
 PH Peptide 1..35
 FT /label= Sig_peptide
 FT Protein 36..302
 FT /label= A-domain

Sequence 576 AA:

```

Query Match      99.1%; Score 1010; DB 20; Length 576;
Best Local Similarity 100.0%; Pred. No. 1.5e-98;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PKQYPIINFITAGATVQSYTNFIRAVRGRLLTGADVRHEIPVLPNRVGLPINQRFILVEL 62
Db 38 PKQYPIINFITAGATVQSYTNFIRAVRGRLLTGADVRHEIPVLPNRVGLPINQRFILVEL 97
QY 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFGGN 122
Db 98 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDNOEDAEATHLFTDVQNRYYTFAFGGN 157
QY 123 YDRLEQLAGNLRENIELGNGLPELEAISALYYSTGGTQLPTLARSFIIQIMISEAARFQ 182
Db 158 YDRLEQLAGNLRENIELGNGLPELEAISALYYSTGGTQLPTLARSFIIQIMISEAARFQ 217
QY 183 YIEGEMETRIRYNRRS 198
Db 218 YIEGEMETRIRYNRRS 233

```

Search completed: February 10, 2004, 16:22:27
Job time : 33.7583 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 16:18:30 ; Search time 11.4485 Seconds
(without alignments)
731.761 Million cell updates/sec

Title: US-10-083-336A-7

Perfect score: 1019

Sequence: 1 MVPKQYPIINFNTAGATVQS.....ARFQYIEGEMRTIRYNRES 198

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCUS COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1019	100.0	267	1	US-08-218-303-16
2	1019	100.0	267	2	US-08-338-793D-61
3	1019	100.0	267	4	US-09-538-873-1
4	1010	99.1	267	1	US-07-901-707-1
5	1010	99.1	267	1	US-07-988-430-1
6	1010	99.1	267	1	US-08-425-336-1
7	1010	99.1	267	1	US-08-488-1138-1
8	1010	99.1	267	1	US-08-477-484B-1
9	1010	99.1	267	2	US-08-646-360-1
10	1010	99.1	267	3	US-08-839-765-1
11	1010	99.1	267	3	US-09-136-389-1
12	1010	99.1	267	4	US-09-610-838-1
13	1010	99.1	267	5	PCT-US92-09487-1
14	1010	99.1	268	2	US-08-356-786-8
15	1010	99.1	290	1	US-08-378-761A-27
16	1010	99.1	290	1	US-08-485-286-27
17	1010	99.1	290	6	5248606-4
18	1010	99.1	534	2	US-08-356-786-10
19	920.5	90.3	540	1	US-08-378-761A-77
20	920.5	90.3	540	1	US-08-485-286-77
21	342	33.6	247	1	US-08-488-1138B-6
22	342	33.6	247	1	US-08-477-484B-6
23	342	33.6	247	2	US-08-646-360-6
24	342	33.6	247	3	US-08-839-765-6
25	342	33.6	247	3	US-09-136-389-6
26	342	33.6	247	4	US-09-610-838-6
27	342	33.6	267	1	US-08-378-761A-74

28 342 33.6 267 1 US-08-485-286-74
29 342 33.6 289 1 US-07-923-692C-4
30 342 33.6 289 1 US-08-184-237-4
31 342 33.6 289 1 US-08-482-920-4
32 342 33.6 289 3 US-08-484-341-4
33 342 33.6 289 3 US-08-483-502-4
34 342 33.6 289 4 US-09-726-651A-4
35 341.5 33.5 282 1 US-08-324-301-15
36 329.5 32.3 250 1 US-08-378-761A-71
37 329.5 32.3 250 1 US-08-485-286-71
38 323.5 31.7 251 4 US-09-538-873-3
39 312.5 30.7 255 1 US-07-901-707-6
40 312.5 30.7 255 1 US-07-988-430-6
41 312.5 30.7 255 5 PCT-US92-09487-6
42 312.5 30.6 248 3 US-08-902-486-7
43 312 30.6 290 1 US-08-245-754A-2
44 312 30.6 290 2 US-08-597-731-2
45 312 30.6 290 2

ALIGNMENTS

RESULT 1

US-08-218-303-16
; Sequence 16, Application US/08218303
; Patent No. 5547867
; GENERAL INFORMATION:
; APPLICANT: Kara, Bhupendra V.
; APPLICANT: Hockney, Robert C.
; APPLICANT: Fittton, John E.
; TITLE OF INVENTION: FERMENTATION PROCESS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman, Darby & Cushman
; STREET: 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,303
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/841,533
; FILING DATE: 26-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Kokulis, Paul N.
; REGISTRATION NUMBER: 16,773
; REFERENCE/DOCKET NUMBER: PNK/3893/94908/MJW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 267 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-218-303-16

Query Match 100.0%; Score 1019; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.3e-111;
Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVPKQYPIINFNTAGATVQS...YTFIRAVRGLITGADVREIPLVLPNVRGLPINQRFILV 60

Db 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNVRGLPINQRFILV 60
 QY 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 120
 Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 120
 QY 121 GNYDRLEQLAGNLRENIELGNGLPLEEAIISALYYSTGGTQPLTLARSLFIICQMISEAAR 180
 Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEAIISALYYSTGGTQPLTLARSLFIICQMISEAAR 180
 QY 181 FOYIEGEMTRIRYNRRS 198
 Db 181 FOYIEGEMTRIRYNRRS 198

RESULT 2

US-08-338-793D-61
 ; Sequence 61, Application US/08338793D
 ; Patent No. 5840521
 ; GENERAL INFORMATION:
 ; APPLICANT: Barth, Peter Thomas
 ; TITLE OF INVENTION: VECTOR
 ; NUMBER OF SEQUENCES: 61
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CUSHMAN DABBY CUSHMAN
 ; ADDRESSEE: INTELLECTUAL PROPERTY GROUP OF
 ; ADDRESSEE: PILLSBURY MADISON & SUTRO, L.L.P.
 ; STREET: 1100 New York Avenue, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20005-3918
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
 ; COMPUTER: IBM PC/XT/AT Compatibles
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: Microsoft Word or ASCII editors
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/338,793D
 ; FILING DATE: 08-Nov-94
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/842,081
 ; FILING DATE: 26-Feb-92
 ; CLASSIFICATION: 435
 ; APPLICATION NUMBER: 9104017.0
 ; FILING DATE: 26-Feb-91
 ; APPLICATION NUMBER: 9109188.4
 ; FILING DATE: 29-Apr-91
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kokulis, Paul N.
 ; REGISTRATION NUMBER: 16,773
 ; REFERENCE/DOCKET NUMBER: DJB/9901/215431/TGW
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-861-3000
 ; TELEFAX: 202-822-0944
 ; TELEX: 6714627 CUSH
 ; INFORMATION FOR SEQ ID NO: 61:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 267 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: Single
 ; TOPOLOGY: Linear
 ; US-08-338-793D-61

Query Match 100.0%; Score 1019; DB 2; Length 267;
 Best Local Similarity 100.0%; Pred. No. 1.3e-111;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNVRGLPINQRFILV 60
 Db 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNVRGLPINQRFILV 60

QY 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 120
 Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 120
 QY 121 GNYDRLEQLAGNLRENIELGNGLPLEEAIISALYYSTGGTQPLTLARSLFIICQMISEAAR 180
 Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEAIISALYYSTGGTQPLTLARSLFIICQMISEAAR 180
 QY 181 FOYIEGEMTRIRYNRRS 198
 Db 181 FOYIEGEMTRIRYNRRS 198

RESULT 3

US-09-538-873-1
 ; Sequence 1, Application US/09538873
 ; Patent No. 6566500
 ; GENERAL INFORMATION:
 ; APPLICANT: VITETTA, ELLEN S.
 ; APPLICANT: GHETIE, VICTOR F.
 ; APPLICANT: SMALLSHAW, JOAN
 ; APPLICANT: BALUNA, ROXANA G.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR AMELIORATING VASCULAR LEAK
 ; FILE REFERENCE: UTSD:603
 ; CURRENT APPLICATION NUMBER: US/09/538,873
 ; CURRENT FILING DATE: 2000-03-30
 ; EARLIER APPLICATION NUMBER: 60/126,826
 ; EARLIER FILING DATE: 1999-03-30
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 267
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
 ; OTHER INFORMATION: Peptide
 ; US-09-538-873-1

Query Match 100.0%; Score 1019; DB 4; Length 267;
 Best Local Similarity 100.0%; Pred. No. 1.3e-111;
 Matches 198; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNVRGLPINQRFILV 60
 Db 1 MYPKQPIINFTTAGATVQSYTNFIRAVGRLLTGADVRHEIPVLPNVRGLPINQRFILV 60
 QY 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 120
 Db 61 ELSNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAIHTLFTDVQNRYYTFAFG 120
 QY 121 GNYDRLEQLAGNLRENIELGNGLPLEEAIISALYYSTGGTQPLTLARSLFIICQMISEAAR 180
 Db 121 GNYDRLEQLAGNLRENIELGNGLPLEEAIISALYYSTGGTQPLTLARSLFIICQMISEAAR 180
 QY 181 FOYIEGEMTRIRYNRRS 198
 Db 181 FOYIEGEMTRIRYNRRS 198

RESULT 4

US-07-901-707-1
 ; Sequence 1, Application US/07901707
 ; Patent No. 5376546
 ; GENERAL INFORMATION:
 ; APPLICANT: Bernhard, Susan L.
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Steve F.
 ; APPLICANT: Lane, Julie A.
 ; TITLE OF INVENTION: Materials Comprising and Methods of
 ; TITLE OF INVENTION: Composition and Use for Ribosome-Inactivating Proteins

; NUMBER OF SEQUENCES: 57
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; ADDRESSEE: Bicknell
 ; STREET: Two First National Plaza, 20 South Clark
 ; STREET: Street
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60603
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/901,707
 ; FILING DATE: 19920619
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/787,567
 ; FILING DATE: 04-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 537654and, Greta E.
 ; REGISTRATION NUMBER: 35,302
 ; REFERENCE/DOCKET NUMBER: 27129/30910
 ; TELEPHONE: (312) 346-5750
 ; TELEFAX: (312) 984-5750
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 267 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ;
 ; US-07-901-707-1

Query Match 99.1%; Score 1010; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 1.5e-110;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILVEL 62
 Db 3 PKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILVEL 62
 QY 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVTFAGGN 122
 Db 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVTFAGGN 122
 QY 123 YDRLEQLAGNLRNIEIENGNGPLLEAISALYYSTGGTQLPTLARSFICIQMISEAARFQ 182
 Db 123 YDRLEQLAGNLRNIEIENGNGPLLEAISALYYSTGGTQLPTLARSFICIQMISEAARFQ 182
 QY 183 YIEGEMTRIRYNRRS 198
 Db 183 YIEGEMTRIRYNRRS 198

RESULT 5
 US-07-988-430-1
 ; Sequence 1, Application US/07988430
 ; Patent No. 5416202
 ; GENERAL INFORMATION:
 ; APPLICANT: Bernhard, Susan L.
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Stephen F.
 ; APPLICANT: Lane, Julie A.
 ; APPLICANT: Lei, Shau-Ping
 ; TITLE OF INVENTION: Materials Comprising and Methods of
 ; PREPARATION AND USE FOR RIBOSOME-INACTIVATING PROTEINS
 ; NUMBER OF SEQUENCES: 101
 ; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
 ; ADDRESSEE: Bicknell
 ; STREET: Two First National Plaza, 20 South Clark
 ; STREET: Street
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60603
 ;
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/988,430
 ; FILING DATE: 19921209
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/901,707
 ; FILING DATE: 19-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/787,567
 ; FILING DATE: 04-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5416202and, Greta E.
 ; REGISTRATION NUMBER: 35302
 ; REFERENCE/DOCKET NUMBER: 31133
 ; TELEPHONE: (312) 346-5750
 ; TELEFAX: (312) 984-9740
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 267 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ;
 ; US-07-988-430-1

Query Match 99.1%; Score 1010; DB 1; Length 267;
 Best Local Similarity 100.0%; Pred. No. 1.5e-110;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILVEL 62
 Db 3 PKQYPIINFTAGATVQSYTNFIRAVRGRLTTGADVRRHEIPVLPNVRVGLPINQRFILVEL 62
 QY 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVTFAGGN 122
 Db 63 SNHAELSVTLALDVTNAYVVGVRAGNSAYFFHPDQEDAEAI THLFTDVQNRVTFAGGN 122
 QY 123 YDRLEQLAGNLRNIEIENGNGPLLEAISALYYSTGGTQLPTLARSFICIQMISEAARFQ 182
 Db 123 YDRLEQLAGNLRNIEIENGNGPLLEAISALYYSTGGTQLPTLARSFICIQMISEAARFQ 182
 QY 183 YIEGEMTRIRYNRRS 198
 Db 183 YIEGEMTRIRYNRRS 198

RESULT 6
 US-08-425-336-1
 ; Sequence 1, Application US/08425336
 ; Patent No. 5621083
 ; GENERAL INFORMATION:
 ; APPLICANT: Better, Marc D.
 ; APPLICANT: Carroll, Stephen F.
 ; APPLICANT: Studnika, Gary M.
 ; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 ; PROTEINS
 ; NUMBER OF SEQUENCES: 140
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606-6402

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/425,336
FILING DATE: 18-APR-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/064,691

FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: Meyers, Thomas C.

REGISTRATION NUMBER: P-36,989

REFERENCE/DOCKET NUMBER: 31394

TELEPHONE: 312/474-6300

TELEFAX: 312/474-0448

TELEX: 25-3856

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-425-336-1

Query Match 99.1%; Score 1010; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.5e-110;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3	PKOYPIINF	TAGATVQSYN	IFRAVRGR	LTGADV	RHEIPV	LPNVRG	LPI	NQRFIL	VEL 62
Db	3	PKOYPIINF	TAGATVQSYN	IFRAVRGR	LTGADV	RHEIPV	LPNVRG	LPI	NQRFIL	VEL 62
QY	63	SNHAELSV	TLALDVT	NAYVVG	YRAGNS	AYFFH	PDNQ	EDAE	AI	THLFTD
Db	63	SNHAELSV	TLALDVT	NAYVVG	YRAGNS	AYFFH	PDNQ	EDAE	AI	THLFTD
QY	123	YDRLEQ	LAGN	REN	IEL	GNGL	PLEE	AI	SAL	YVY
Db	123	YDRLEQ	LAGN	REN	IEL	GNGL	PLEE	AI	SAL	YVY
QY	183	YIEGEM	TRIR	YNRS	198					
Db	183	YIEGEM	TRIR	YNRS	198					

RESULT 7
US-08-488-113B-1
Sequence 1, Application US/08488113B
Patent No. 5744580
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Stephen F.
APPLICANT: Studnika, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/488,113B
FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/425,336

FILING DATE: 18-APR-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 11022US07/200-70.P3.C2A

TELEPHONE: 312/707-9155

TELEFAX: 312/707-8889

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 267 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-488-113B-1

Query Match 99.1%; Score 1010; DB 1; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.5e-110;
Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3	PKOYPIINF	TAGATVQSYN	IFRAVRGR	LTGADV	RHEIPV	LPNVRG	LPI	NQRFIL	VEL 62
Db	3	PKOYPIINF	TAGATVQSYN	IFRAVRGR	LTGADV	RHEIPV	LPNVRG	LPI	NQRFIL	VEL 62
QY	63	SNHAELSV	TLALDVT	NAYVVG	YRAGNS	AYFFH	PDNQ	EDAE	AI	THLFTD
Db	63	SNHAELSV	TLALDVT	NAYVVG	YRAGNS	AYFFH	PDNQ	EDAE	AI	THLFTD
QY	123	YDRLEQ	LAGN	REN	IEL	GNGL	PLEE	AI	SAL	YVY
Db	123	YDRLEQ	LAGN	REN	IEL	GNGL	PLEE	AI	SAL	YVY
QY	183	YIEGEM	TRIR	YNRS	198					
Db	183	YIEGEM	TRIR	YNRS	198					

RESULT 8
US-08-477-484B-1
Sequence 1, Application US/08477484B
Patent No. 5756699
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
APPLICANT: Carroll, Stephen F.
APPLICANT: Studnika, Gary M.